the RESEARCH Quarterly

of the American Association for Health,
Physical Education, and Recreation,
Department of the National Education Association,
1201 Sixteenth St., N.W., Washington 6, D. C.

Vol. 25, No. 2 **MAY 1954**

CONTENTS

Published in March, May, October, and December by the American Association for Health, Physical Education, and Recreation, 1201 Sixteenth Street, N.W., Washington 6, D. C. Subscription available to libraries, institutions, and departments at $5.00 per year. Of the Professional Membership ($10.00 per year), $3.00 covers subscription to the *Quarterly.* Of the Student Professional Membership ($5.00 per year), $1.50 is for subscription to the *Quarterly. Subscriptions to individuals through membership only.* Single copies $1.25.

Send membership dues or subscriptions to Circulation Dept., 1201 Sixteenth Street, N.W., Washington 6, D. C. Editorial Office, 1201 Sixteenth Street, N.W., Washington 6, D. C. Entered as second-class matter at the Post Office at Washington, D. C., under the act of March 3, 1879. Additional entry at Baltimore, Md.

Selection of Men Teachers for Positions in High School Physical Education

AUBREY A. BATES
Sacramento State College
Sacramento, California

Abstract

This article summarizes practices and recommendations contained in questionnaires from 383 physical educators, 224 high school district superintendents, and 96 directors of teacher placement agencies. Qualifying standards and administrative principles currently used in selecting secondary physical education teachers are identified, and pertinent findings are cited from related research studies.

Sources and kinds of information used to evaluate an applicant's desirability for appointment were ranked in significance as follows by the questionnaire respondents: (1) personal interview; (2) observation in a teaching situation; (3) reference letters; (4) college transcripts; (5) record of participation in sports; (6) statement of philosophy of physical education; (7) professional and civic activities; (8) written exams; and (9) performance of athletic skills.

TEACHER PERSONNEL in physical education must be competent, energetic, and forceful if high school youth, through their physical education experiences, are to acquire the skills, understandings, and appreciations requisite to wholesome growth and adjustment. Many of the currently recognized leaders in physical education believe the most pressing problem of the profession lies in the need for an upgrading in the quality of teacher personnel for this curricular area.

The Problem

The purpose in this study[1] was to formulate a statement of policies, practices, and devices whereby the eligibility ranking of candidates for a position in secondary physical education for boys could be determined, and wherein the derived teacher selection program would reflect (1) the recommendations of physical educators vitally concerned with the status of the profession, (2) the practices relied upon by school superintendents charged with the selection of competent teachers, and (3) the observations and opinions of directors of teacher placement agencies.

Method of Procedure

The normative survey method was utilized in conducting the study. Survey forms were sent to: (1) 837 graduate members of Phi Epsilon Kappa, national fraternity in physical education; (2) 412 superintendents of selected high school districts, and (3) 254 directors in charge of the member agencies of the National Institutional Teacher Placement Association and the National Asso-

---

[1]This study was made in partial fulfillment of the requirements for a Doctor's degree at the University of California, Los Angeles, 1952.

ciation of Teachers Agencies. Of the total of 1,494 mailed questionnaires, usable returns from 703 were incorporated in machine tabulation of replies.

Review of Related Literature

From the review of selected literature on the selection, appointment, in-service evaluation, and retention of teachers, the following findings and recommendations seem most pertinent to the problem:

1. Candidates for certification as qualified teachers of secondary physical education, in addition to meeting the requirements applicable to all teachers in general, should meet these stipulations: (a) be able to participate in collegiate competitive sports in terms of health, physique, and temperament; (b) possess personal qualities requisite to effective and wholesome leadership of adolescent boys; and (c) demonstrate attainments in poise, speech, appearance, and intelligence deemed essential to meet community, social, and civic responsibilities and to cope with public relations problems normally to be encountered (7).

2. The 120 units, or four years of college education, required for the certification of secondary physical education teachers, should include approximately 20 units in professional education, 25 units in foundation sciences, 30 units in general education, and 40 units in physical education (2).

3. Teacher-education institutions, in their programs for physical education majors, should place greater emphasis upon leadership of student activities, public relations, health education, teaching in rural schools, and over-all improvement in scholarship (1).

4. The sources most frequently used to gain information about prospective applicants were: (a) formal application blank; (b) transcripts of educational experiences, and credentials; (c) reference letters from former employers and professional associates; (d) personal interviews; (e) observations in teaching situations; (f) health examinations; (g) written and oral examinations; (h) statements of educational philosophy and goals; and (i) autobiographical listing of cultural, social, and political affiliations and experiences (5).

5. All information used to appraise an applicant's qualifications should come from authentic sources, be accurate and complete in content, and be obtained under conditions compatible with ethical standards of the profession (3).

6. The attributes and qualities most characteristic of excellent teachers were: (a) skills in instruction; (b) discipline and control of students; (c) professional knowledge in subject area taught; (d) desirable personality traits and (e) desirable character traits (4).

7. The traits and qualities most frequently considered by administrators when selecting secondary physical education teachers were: (a) personality, (b) character, (c) instruction, (d) knowledge, and (e) discipline (7, p. 112-113).

8. Aspects of teacher selection programs frequently shown to be contributors to inefficiency were: (a) use of invalid indices of teaching success; (b)

incompetency of personnel who made evaluations of an applicant's potentialities; and (c) faulty procedures and policies for contracting, interviewing, and appointing teacher candidates (6).

Findings of the Study

Data obtained in the questionnaires from the 383 physical educators, 224 superintendents, and 96 placement directors, contained this information:

1. The superintendents and placement directors ranked the SOURCES AND KINDS OF INFORMATION used to assess the qualifications of men applicants as follows: (a) personal interview; (b) observation of teaching; (c) reference letters from professional associates; (d) transcript of college experiences; (e) participation in sports; (f) statement of philosophy; (g) professional and civic activities; (h) written examinations; and (i) performance test and athletic skills (See Table 1).

2. In obtaining the names of qualified applicants, the superintendents relied almost entirely upon recommendations solicited from teacher-education

TABLE 1

Weighted Importance of Various Kinds of Information About Men Applicants for Positions in High-School Physical Education

Kind of Information	Physical Educators N	Physical Educators Mean	Superintendents N	Superintendents Mean	Placement Directors N	Placement Directors Mean
Personal interview with the superintendent, principal, or other designated interviewer	379	2.73	219	3.28	93	3.53
Observation and rating in teaching situations	379	3.21	204	2.94	91	3.08
Information obtained from references listed by applicants	373	2.09	214	2.72	93	3.01
Transcripts of college courses and grades	374	2.40	214	2.60	92	2.21
Record of participation in college and professional sports	377	1.38	206	1.69	91	2.51
Applicant's written statement of his philosophy of physical education	374	1.94	186	1.79	82	2.29
Membership and activities in professional and civic organizations	372	1.91	201	1.57	88	1.42
Score on a written examination	376	1.70	146	1.60	70	.80
Score on a performance test in physical education skills	373	2.05	147	1.04	71	1.02
Average Weighting		2.61		2.14		2.10

NOTE: This table should be read as follows: the mean importance weighting of information obtained about applicants through the personal interview was 2.73 by th 379 physical educators; 3.28 for the 219 superintendents; and 3.53 for the 93 placement directors. Higher figures for the mean denote greater importance of factor.

institutions. Also in areas where teacher shortages existed, school boards or superintendents frequently utilized the services of commercial placement agencies to secure the names of prospective applicants.

3. In 37.3 per cent of the 224 high school districts, MEN WERE SELECTED TO FILL VACANT POSITIONS in high school physical education through agreement between the district superintendent and the principal of the high school in which the vacancy existed.

4. SALARIES PAID TO HIGH SCHOOL PHYSICAL EDUCATION TEACHERS were slightly higher than those paid to teachers of other subjects. Most of the school districts granted additional compensation to coaches for the performance of extra duties relative to after school competitive athletics.

5. Relative to QUALIFYING STANDARDS OR CONDITIONS WHEREBY THE LESS DESIRABLE APPLICANTS CAN BE IDENTIFIED, most employing school districts apparently follow the policy of classifying as "preferred" only those applicants who rate satisfactorily in regard to all screening factors.

The following recommendations reflect current practices: (a) PROFESSIONAL EDUCATION. The applicant should possess an A.B. degree with a teaching major in physical education; (b) PHYSICAL HANDICAPS. The applicant should have no physical handicap extensive enough to reduce efficiency in the performance of specified duties; (c) ATHLETIC RECORD. Slight preference should be given to applicants who have outstanding athletic records; (d) AGE. The applicant's age should not be a detriment to the efficient performance of his duties; (e) HEALTH. Certification by a licensed physician that the applicant's health is satisfactory for teaching should be required; (f) EXPERIENCE. Qualifications for appointment should not require prior teaching experience, although credit should be allowed to applicants who have had satisfactory teaching experience; (g) RESIDENCE. The place of residence should not influence his rating or qualification for appointment.

6. The superintendent and placement directors indicated the following objectives of the PERSONAL INTERVIEW, in rank order of importance: (a) to appraise the applicant's appearance, professional poise, and personality; (b) to ascertain the applicant's basic philosophy of physical education; (c) to estimate the applicant's knowledge of physical education; (d) to discuss probable placement, salary, and contract; (e) to ascertain the applicant's general educational knowledge; and (f) to learn of the applicant's future ambitions and plans.

7. REFERENCE LETTERS should disclose information about the applicant in terms of his ability to fulfill the duties of the position sought, his character, personality, and professional attitude, and should be solicited from a person professionally associated with the applicant.

8. Aspects of COLLEGE TRANSCRIPTS that bear significantly upon the applicant's likelihood of success in teaching include: (a) number and kind of courses completed in physical education; (b) status of institution sending transcript; (c) units and grades in teaching major; (d) credentials and

degrees obtained; and (e) record of participation in special activities, including sports.

9. A majority of the respondents in each of the three groups did not favor the use of WRITTEN EXAMINATIONS as an appraisal device.

10. INEFFICIENCY IN TEACHER SELECTION apparently is influenced materially by many factors other than the incompetency of candidates. Five of such negative factors or obstacles are identified as follows: (a) OVEREMPHASIS ON WINNING TEAMS (66%).[2] "Local community pressure groups demand that the physical education teacher be selected primarily for his ability to produce winning athletic teams"; (b) SALARIES NOT COMPATIBLE WITH CALIBER OF MEN NEEDED (49.1%). "Salaries are too low to attract capable men"; (c) MANY GOOD PHYSICAL EDUCATION TEACHERS TRANSFER TO LESS PUBLICIZED SUBJECT AREAS (43.9%), "Community criticism and adverse publicity which accrue to poor athletic teams discourage some qualified men from electing physical education as the major subject area in which to teach"; (d) EMPLOYMENT PRACTICES FREQUENTLY REQUIRE THAT PHYSICAL EDUCATION TEACHERS MUST TEACH OTHER SUBJECTS IN ADDITION TO PHYSICAL EDUCATION (42.5%), "Too many physical education applicants are not qualified to teach subjects other than physical education"; (e) LARGE CLASSES, POOR FACILITIES, LONG HOURS DISCOURAGE SOME GOOD MEN (37.3%), "Teaching conditions discourage some qualified men from electing physical education as the major subject area in which to teach."

Administrative Principles for the Teacher Selection Program—Data of the study identified the following policies and practices as minimal to an efficient teacher selection program:

1. DEFINE THE JOB. Vacant positions should be adequately defined in terms of teaching duties to be performed, qualifying credentials, responsibilities in conducting extracurricular activities including competitive sports, estimated number of duty hours per week, expected participation in school and community activities, salary to be received including expenses and pay for extra duty, and opportunity for tenure and advancement.

2. SECURE QUALIFIED APPLICANTS. Superintendents should utilize as many of the sources for securing the names of prospective applicants as needed to insure the availability of a well-qualified teacher to fill existing vacancies.

3. APPRAISE THE MERITS OF APPLICANTS. The selection of teachers should be based upon factors found by research to be contributive to teaching success. Factors currently used should be evaluated in terms of their validity.

4. DERIVE APPRAISALS OBJECTIVELY. An evaluation committee of not fewer than three competent persons should independently appraise the qualifications of an applicant, and the average rating by the total committee should be used as the final index of the applicant's eligibility.

5. APPOINT THE BEST AVAILABLE CANDIDATE. Eligibility lists for specified subject areas should be established and utilized by superintendents. The

---

[2]Mean of the three groups.

superintendent and school principal should nominate an applicant from those ranking highest on the eligibility list, and the board of education should appoint the nominee.

6. Continue the Evaluation of the Appointee in Service. Selection and appointment policies used by employing school systems should stipulate a probationary period of teaching during which further evaluations of the candidate's qualifications should be accomplished to determine whether or not the teacher is deserving of advancement to tenure status.

Conclusions and Comments

In terms of the major purposes and in view of the basic findings of the investigation, these conclusions and comments seem justified by the data:

1. Selection Practices May Be Faulty. Teacher selection and appointment policies frequently are neither sufficiently professional in nature nor adequately separated from the community politics. Personal friendships between applicants and local school officials, status ambitions of school administrators, commercial interests of community groups, and desires for public esteem by school board members are factors which sometimes cause the selection of a physical education teacher who later becomes a liability rather than an asset to the best interests of the school and community.

2. Participation in Sports Seems Advantageous to Teaching. Participation in major competitive sports is believed by both administrators and teachers to constitute an integral part of the applicant's background of skills and experiences essential to competency in this curricular area. The identification of isolated instances of men who became successful physical education teachers without having had such sports experiences does not adequately refute the claim that their success possibly would have been greater had they participated in college athletics.

3. Greater Standardization and Clarity in College Transcripts Seems Needed By Teacher Selection Agencies. The lack of uniformity in college transcripts seemingly adds much to the complexity of the task of transposing the factual data contained in such transcripts to any desired index of the applicant's competency as a teacher. It apparently would be a great service to teacher selection agencies if the transcripts of candidates for teaching credentials were standardized as to format, symbols used, and cumulative summaries of data.

4. Coaches of Athletic Teams Also Should Be Competent Teachers. Apparently the administrator's philosophy of physical education materially influences the quality of the program conducted by the physical education teachers in the school. For example, "pressure" upon the coach that he produce winning teams probably represents the most identifiable cause for overemphasis upon competitive athletics. Administrative "guidance" should cause all phases of the program to be conducted properly, and should ensure that only good coaches who also are good teachers will be rated as satisfactory for selection or retention.

5. EVERY SCHOOL SYSTEM SHOULD ATTEMPT TO BE EFFICIENT IN SELECTING TEACHERS. An efficient program of selecting teachers for positions in secondary school physical education, even for small rural schools, should include the following minimum essentials: (a) encouragement of several qualified candidates to apply for each vacant position; (b) the processing of all applications in an objective and uniform manner covering at least an evalution of college trnscripts, reference letters, and the personal interview; and (c) appointment of the highest ranking applicant.

6. TEACHER SELECTION COMMITTEES SHOULD BE COMPETENT. Personnel of the teacher selection committee should be chosen carefully in terms of their willingness, availability, and competency in deriving impartial and reliable ratings for all applicants.

7. THE POLICY OF GRANTING EXTRA COMPENSATION TO ATHLETIC COACHES NEEDS CRITICAL EVALUATION. The growing trend to grant extra compensation to physical education teachers for the extra time devoted to competitive athletics should be studied carefully in terms of the implications of such policies to the total teaching profession. It seems unwise to show privilege to any teacher group, and there are many who contend that physical education teachers who receive extra pay are being shown undue favoritism. Perhaps a more wholesome solution to the problem would stem from a re-study and redistribution of teacher load for all teachers based upon the expenditure of energy and time, and upon duties and responsibilities carried out.

8. PERSONNEL ADMINISTRATION. Differences in practices among varying sections of the nation point to the need for research to determine if a better instructional program in physical education results when separate personnel are employed to coach competitive sports from those selected to conduct the regular physical education program.

REFERENCES

1. BENDER, PAUL F. *A Study of Some Factors Affecting the Appointment of Prospective Teachers of Physical Education* (unpublished doctoral dissertation, New York University, 1939), p. 121.

2. BLESH, T. ERWIN, "Evaluative Criteria in Physical Education," *Research Quarterly*, XVII (May 1946), 115-16.

3. COULBOURN, JOHN, *Selection of Teachers in Large City School Systems*, Contributions to Education, No. 470 (New York: Columbia University, 1948), pp. 79-80.

4. DANIEL, JOSEPH MCTYEIRE, *Excellent Teachers: Their Qualities and Qualifications*, (Columbia: University of South Carolina, 1944), p. 308.

5. DEFFENBAUGH, W. S., AND WILLIAM H. ZIEGEL, JR., *Selection and Appointment of Teachers*, Bulletin, 1932, no. 17, National Survey of Secondary Education, Monograph no. 12 (Washington, D. C.: U. S. Government Printing Office, 1933), pp. 47-74.

6. LONG, CHARLES MILTON, *Ten Undesirable Criteria often Used in the Selection and Retention of Teachers*, (unpublished master's thesis, Colorado State College of Education, 1938), pp. 51-83.

7. MINNEGAN, DONALD I., *Criteria of a good Physical Education Program for Boys in the Senior High School* (unpublished doctoral dissertation, George Washington University, Washington, D. C., 1947), p. 411.

(Submitted 5-8-53)

Relationship of Strength and Anthropometric Measures to Various Arm Strength Criteria

H. HARRISON CLARKE
University of Oregon
Eugene, Oregon

Abstract

This article reports the results of three studies on various strength and anthropometric factors entering dominantly into co-ordinated movements of the arm and shoulder girdle area. Best results were obtained with arm strength formulae based on the number of pull-ups and push-ups performed by the subjects. The factors with highest multiple correlational relationships with these formulae were girth of flexed-tensed upper arm, standing height, and strength of the shoulder inward rotator muscles.

FOR SEVERAL YEARS, studies were conducted in the Springfield College Physical Education Research Laboratory[1] on various factors entering dominantly into co-ordinated strength movements of the arm and shoulder girdle area. In this process, nine tests of muscular strength and five anthropometric measurements were related to each other and to 12 arm strength criteria.

The Experimental Variables

STRENGTH TESTS

Strength tests of the following nine muscle groups were utilized in the studies: elbow extensors and flexors; and shoulder extensors, flexors, adductors, abductors, outward rotators, inward rotators, and shoulder horizontal adductors. Strength was measured with a cable tensiometer applied to a specially designed pulling assembly, in accordance with techniques described by Clarke (1). The testing was unilateral only, the left arm and shoulder girdle being used.

ANTHROPOMETRIC TESTS

Five anthropometric tests were used, as follows: girths of upper arm relaxed, flexed, and "tension flexed"; length of upper arm; and standing height. With the exception of standing height, the measuring instrument was a cloth tape with a Gulick handle. The spring located in this handle permitted the application of uniform tension, which was of special value in making objective girth measurements (7). Coefficients of correlation obtained in this study for the various anthropometric measures by the test-retest method utilizing different testers were between .90 and .95. The testing techniques were as follows:

[1]Conducted with the assistance of many physical education graduate students at Springfield College, in part under subsidy from the Office of Naval Research. Special acknowledgments are made to Theodore L. Bailey and Donald K. Matthews, laboratory supervisors during the period of this experimentation.

1. *Standing Height.* Subject stood erect in stocking feet on platform of weight balance scale, back to measuring rod; height recorded to nearest inch.

2. *Length of Upper Arm.* Subject sat erect in chair, feet on floor; arm adducted at shoulder to 90°, forearm flexed at elbow to 90° palm downward, fingers extended; acromion process and lateral epicondyle located by palpation and marked with a flesh pencil; distance measured between pencil marks with tape.

3. *Girth of Upper Arm "Tensed Flexed."* Subject sat erect in chair, feet on floor; arm flexed at shoulder to 90°, elbow flexed with forearm in supine position; fist clenched and biceps contracted to form as large a muscle as possible; girth measured around most prominent part of upper arm, bringing the pressure on the tape each time to the same point on the spring in the Gulick handle. (Techniques for testing the girths of the upper arm flexed and relaxed are not given, as these measures did not appear prominently in the results of the study.)

Arm Strength Criteria

For purposes of analysis, the following 12 composite scores, indexes, and tests were used individually as criteria: composites (raw-score and Hull-score), indexes (McCloy's, Rogers', various pull-up and push-up combinations with weight), and tests (MacCurdy's, pull-ups, push-ups).

COMPOSITE STRENGTH SCORES

Raw-score Composite. For this criterion, the raw scores obtained by each subject on the nine strength tests were added. This method had the effect of weighting the tests, as small muscle movements with low strength scores were combined directly with large movements with high scores. Thus, the weight in the criterion for each of the tests was proportionate to the actual strength of the movement.

Hull-Score Composite. The Hull-score values obtained for each subject on the nine strength tests were added to form this criterion. Hull scores are based on the standard deviation distances of scores from their respective means (a range of three and one-half S.D.'s either side of the mean), and the use of this scale has the effect of giving each test equal weight in the criterion.

ARM STRENGTH FORMULAE

Rogers' Arm Strength Score. This arm strength score, a part of the Rogers Strength Index battery, is based upon the number of pull-ups (from chinning bar) and push-ups (from parallel bars) the subjects can perform. The pull-up and push-up tests were administered as described by Clarke (2). The arm strength score was computed from the following formula:

$$\text{A.S.} = (\text{Push-ups} + \text{Pull-ups})\ (\text{Wgt.}/10 + \text{Hgt.} - 60)$$

McCloy's Arm Strength Score: The same pull-ups and push-ups used for the Rogers' score were utilized in the McCloy (6) formula:

$$\text{A.S.} = 1.77\ \text{Wgt.} + 3.42\ (\text{Pull-ups}) - 46$$

The same formula was used for push-ups; and combined pull-up and push-up strength was obtained by adding both.

OTHER PULL-UP AND PUSH-UP ARRANGEMENTS

The number of pull-ups and push-ups were also utilized to form the following five additional arm strength criteria: number of pull-ups, number of push-ups, number of pull-ups times weight, number of push-ups times weight, and number of pull-ups plus push-ups times weight.

MACCURDY'S ARM STRENGTH (PULLING) TEST

Unlike most of the other criteria, the MacCurdy arm strength tests involve direct pulls and pushes on a dynamometer. In this study, the pulling test only was used. The testing techniques are described in MacCurdy's Columbia University doctoral dissertation (5).

Research Procedures

SUBJECTS

Data were collected on 62 Springfield College male students. Inasmuch as most of the subjects were physical education majors, they were thought to be physically superior to college students in general. Believing that the nature of the subjects might be a factor in the results of an experiment requiring vigorous physical effort, data were collected to define further the physical status of the group. As the individual subjects participating in the research varied from year to year, a general summary only of these data is given.

Age. Age was limited to students between 20 and 26 years inclusive; the median was 22.25.

Height. The median height was 69.25 inches; the interquartile range, from 67 to 71 inches; the range, from 64 to 74 inches.

Weight. The median weight was 156 pounds; the interquartile range from 149 to 171 pounds; the range, from 126 to 236 pounds.

Strength Index (2). The median S. I. was 2984 points; the inter-quartile range, from 2688 to 3150 points; the range, from 2000 to 3900 points.

Physical Fitness Index (2). The median PFI was 105; the interquartile range, from 93 to 118; the range, from 56 to 150.

Somatotype (8). The subjects were dominantly mesomorphic, with a pronounced drift toward the meso-ectomorph.

TRAINING OF TESTERS

Before the testers were permitted to record test scores of either the experimental variables or the criterion measures, they were required to demonstrate testing competence, as follows: (a) They were first instructed in the testing techniques and practiced testing with available subjects. (b) Then, the testers administered each test twice to the 62 subjects and correlated the results to obtain an objectivity coefficient. This process was continued until the testers achieved agreement equal to the objectivity coefficients reported for each test. Approximately three weeks of constant testing were necessary for the testers to reach this level of competence for all tests.

SUBJECT FATIGUE

The fatigue element from repeated maximal exertions of the arm and shoulder girdle muscles, while possibly present, was minimized by spaced testing and was equalized by systematically changing the sequence of tests for each testing period.

STATISTICAL ANALYSIS

All experimental variables were intercorrelated and were correlated with each of the arm strength criteria by means of the Pearson product-moment method. In interpreting the correlations, *r's* of .250 and .325 are significant at the .05 and .01 levels of confidence respectively for 62 cases (4).

Coefficients of multiple correlation were computed to select the experimental variables which best represent or account for each of the criterion measures. These coefficients were calculated by the Wherry-Doolittle method (4), as this method permits an automatic succession of variables to be chosen for the multiple, offers a system of mathematical checks on the accuracy of the computation, and allows for stopping the process when R no longer increases appreciably.

Predictive indices were utilized to indicate the percentage better than chance relationship of the various correlations obtained in the study. This statistic also permitted a crude comparison of the relative significance between the various correlation coefficients.

Results of Initial Studies

During the first year of the study, the experimental variables consisted of cable-tension strength tests of nine movements of the elbow and shoulder joints. The arm strength criteria were: McCloy Arm Strength Formula (for pull-ups, push-ups, and pull-ups plus push-ups), Rogers Arm Strength Score, number of pull-ups, number of push-ups, number of pull-ups times weight, number of push-ups times weight, and number of pull-ups plus push-ups times weight.

For the second year, the experimental variables were the same as the first year, with girths of the upper arm relaxed and flexed added. The criterion measures were raw-score composite and Hull-score composite. Significant results of this experimentation are summarized below.

ZERO-ORDER CORRELATIONS

1. The intercorrelations of the cable-tension strength tests were on the whole significant well beyond the .01 level of confidence. Correlations of .61 between shoulder extension and adduction and .59 between shoulder extension and elbow extension were the highest obtained.

2. Girths of the upper arm relaxed and flexed correlated between .28 and .55 with the various cable-tension strength tests. They also correlated .89 with each other, thus demonstrating that only one of these need be included in further research. The flexed position was chosen because of slightly higher correlations with the arm strength criteria.

3. Correlations between the cable-tension strength tests and the criteria composed of the number of pull-ups and push-ups (alone, added, or multiplied by weight) were comparatively low, varying from —.03 between shoulder adduction and number of push-ups to .37 between shoulder inward rotation and push-ups times weight. Twenty-nine—all but seven of these correlations—fell below .25, which is at the .05 level of confidence.

4. In general, the correlations of the experimental variables with the Rogers and McCloy arm strength criteria were higher than those obtained with the other criteria involving the pull-ups and push-ups. These correlations ranged from .06 in the case of shoulder abduction and the Rogers Arm Strength Score and .57 in the case of shoulder extension and McCloy's Arm Strength Score (push-ups). Thirty-three—all but twelve of these correlations—were significant beyond the .05 level.

5. The correlations of the strength test variables were naturally much higher with the raw-score and Hull-score composites. They ranged from .53 between shoulder abduction and Raw-score composite to .81 between shoulder adduction and the same criterion. All of these correlations were significant beyond the .01 level.

MULTIPLE CORRELATIONS

Multiple correlations were computed between the cable-tension strength tests and the various arm strength criteria. The coefficients obtained and the order in which the strength tests appeared in each multiple appear in Table 1. Comments relative to these results follow.

1. The highest multiple correlation was .99, found with each of the composite score criteria. These high correlations were a natural result of the experimental situation, inasmuch as the strength tests, themselves, formed the criterion in each case. The first three tests in the multiple correlation for the raw score composite (shoulder extension, shoulder flexion, and shoulder adduction) had the highest means. Two of the same tests, shoulder adduction and shoulder flexion, also, appeared in the battery for the Hull-score composite.

2. The multiple correlations with the McCloy formulae as the criteria ranked next in size. The highest of these multiples was .66, when push-ups only were used; the other two correlations were .62 and .60. With the Rogers score, R was .52. In terms of predictive index, the McCloy push-up results had 67 per cent more predictive value than did the Rogers results.

3. The lowest multiple correlation of .32 was obtained when the number of pull-ups and push-ups was used separately. Some increase in the multiple was found when these numbers were multiplied by body weight. The multiple of .46, when the number of pull-ups plus push-ups times weight was the criterion, nearly equalled the .52 obtained when the Rogers Arm Score was used.

4. Three strength test items stood out prominently in the number of multiples in which they appeared, as follows:

(a) *Shoulder flexion* appeared in nine of the multiples; was in first or second position in all but one of these.

TABLE 1

Variables Appearing in Multiple Correlations Utilizing Arm Strength Criteria[1]

Strength test variables	Composite raw score	Average Hull score	McCloy pushup formula	McCloy pullup formula	McCloy pullups pushups	Rogers arm strength	PU PU times weight	Pushups times weight	Pullups times weight	Number of pushups	Number of pullups
1. Elbow flexion			2	2	1	1	2		2	2	1
2. Elbow extension		3		3							
3. Shoulder flexion	2	2			2	3	1	1	1	1	2
4. Shoulder extension	1		1								3
5. Shouilder abduction											
6. Shoulder adduction	3	1	3	1	3	4				3	
7. Shoulder horizontal flexion	4										
8. Shoulder outward rotation		4									
9. Shoulder inward rotation					4	2	3	2			
Multiple correlations	.99	.99	.66	.62	.60	.52	.46	.44	.39	.32	.32
Predictive indices	.86	.86	.25	.21	.20	.15	.11	.10	.08	.05	.05

[1]Numbers in columns refer to order in which variables appeared in the multiple correlations.

(b) *Elbow flexion* was found in eight of the batteries; was in either first or second position throughout.

(c) *Shoulder adduction* appeared in seven of the multiple correlations; ranked first in one battery, although its usual position was third or fourth.

5. Two other strength variables should be mentioned: shoulder inward rotation was included in four of the multiple batteries; and shoulder extension appeared first in two batteries. All but one of the other strength variables, shoulder abduction, appeared in one or two multiples, ranking in third or fourth position in each instance.

6. The girth of the upper arm flexed correlated .67 and .63 with raw-score and Hull-score composites respectively.

Latest Study

During the third year of this research, strength and anthropometric variables were selected on the basis of their frequency of appearance in the previous correlations and in their relatively high zero-order correlations with arm strength criteria. Thus, the following tests were included: five strength tests of elbow flexion and shoulder flexion, adduction, inward rotation, and extension; and one anthropometric test of girth of upper arm (changed to "tension flexed," as described above). Two additional anthropometric tests, standing height and length of upper arm, were included among the experimental variables, as crude indicators of the distance each subject lifts himself in pull-ups and push-ups.

For arm strength criteria, those scores with which high multiple correlations had been obtained in the earlier research were included. These were: the McCloy formula for pull-ups, for push-ups, and for pull-ups plus push-ups; and the Rogers Arm Strength Score. The MacCurdy pulling test was added, in order to utilize a criterion based upon the application of the strength of the arm and shoulder girdle muscles in a single muscular effort, rather than by pull-ups and push-ups.

Paired Relationships

The product-moment correlations among the experimental variables and the correlations of the experimental variables with the arm strength criteria appear in Table 2. Observations pertaining to these results appear below.

ANTHROPOMETRIC VARIABLES

The most significant and surprising correlations of the experimental variables with the criteria were .74, .74, and .73 between the girth of tensed-flexed upper arm and the McCloy formulae. The amounts of these correlations have been verified twice in subsequent studies (9). The correlations were higher than any correlation previously obtained against arm strength measures.

The highest correlation of experimental variables with the MacCurdy pulling test was .49, again obtained with the girth measure, although duplicated with the shoulder inward rotation strength test. Correlations of the anthropometric measures with the Rogers Arm Strength Score were so low as to be statistically insignificant.

TABLE 2

Intercorrelations of Experimental Variables and Their Correlations with the Various Arm Strength Criteria

Experimental variables	1	2	3	4	5	6	7	8	C1 Roger's Arm Strength	C2 McCloy's Arm Strength (pullups)	C3 McCloy's Arm Strength (pushups)	C4 McCloy's Arm Strength (pullups and pushups)	C5 MacCurdy (pulling)
1. Shoulder adduction	X	.78	.60	.62	.53	.34	.28	.34	.17	.51	.47	.51	.41
2. Shoulder extension		X	.66	.60	.47	.22	.34	.47	.07	.55	.49	.53	.45
3. Shoulder flexion			X	.56	.47	.19	.11	.52	.17	.53	.56	.59	.41
4. Shoulder inward rotation				X	.48	.29	.19	.41	.22	.51	.53	.58	.49
5. Elbow flexion					X	.18	.41	.42	.29	.39	.48	.49	.38
6. Standing height						X	.54	.07	.14	.34	.35	.25	.46
7. Length of upper arm							X	.30	−.04	.47	.42	.32	.36
8. Girth of flexed-tensed upper arm								X	.04	.74	.74	.73	.49

In general, fairly high correlations were found between the anthropometric measurements and the strength tests. The correlations with the girth of tensed-flexed upper arm were somewhat higher than for the other anthropometric tests. The highest such correlation was .52 with shoulder flexion.

A correlation of .54 was found between standing height and upper arm length. The girth measure had some relationship to upper arm length (.30), but was not significantly related to standing height (.07).

MUSCULAR STRENGTH VARIABLES

A number of strength test variables correlated fairly high with the various arm strength criteria. The highest correlation was .59 between shoulder flexion and McCloy's formula with pull-ups plus push-ups; the lowest, .07 between shoulder extension and the Rogers arm score. All of the strength test correlations with McCloy and MacCurdy criteria exceeded those with the Rogers formula.

The inter-correlations among the strength test variables were quite high, ranging from .47 to .78. As in the initial studies, the highest intercorrelation was between shoulder adduction and shoulder extension. Also, these correlations were consistently higher than in previous reports, owing, it is believed, to greater skill of the testers in administering the tests.

Multiple Correlations

A composite picture of the results of Wherry-Doolittle multiple correlations of the strength and anthropometric test items with the various arm strength criteria appears in Table 3. Comments relative to these results follow:

1. The multiple correlations obtained in the final study were higher than any previously obtained with pull-up and push-up criteria. The highest such

multiple correlation in the initial studies was .66; in this study, three such correlations were around. 80. In all three instances, the McCloy arm strength formulae constituted the criterion measures.

TABLE 3

Strength and Anthropometric Variables Appearing in Multiple Correlations Utilizing Arm Strength Criteria

	Rogers' Arm Strength	McCloy's (pull-ups)	McCloy's (push-ups)	McCloy's (pull-ups and push-ups)	MacCurdy's (pulling)
Anthropometric variables	C1	C2	C3	C4	C5
1. Shoulder adduction	3		3		
2. Shoulder extension					
3. Shoulder flexion			4		
4. Shoulder inward rotation			3	2	3
5. Elbow Flexion	1				
6. Standing height	3	2	2	4	2
7. Length of upper arm	2	4			
8. Girth of flexed-tensed upper arm		1	1	1	1
Multiple *R's*	.36	.81	.8093	.7951	.6655
Predictive indices	.07	.41	.41	.40	.26

2. The multiple correlation with McCloy's push-up formula increased to .87, when elbow and shoulder flexion endurance items were included. These endurance tests were administered in the Kelso-Hellebrandt ergograph in accordance with techniques described by Clarke (3).

3. The next highest multiple correlation was .67, obtained when the MacCurdy pulling test was the criterion. This multiple, also, was higher than those in initial studies.

4. The lowest multiple correlation was .36 with Rogers arm score as the criterion. The earlier correlation with this criterion was .52, with only strength tests used.

5. Considering the variables entering into the multiple correlations, two anthropometric variables proved unexpectedly to be of greatest importance. These were girth of flexed-tensed upper arm and standing height. The girth measure was first in four of the multiples, but did not appear at all when Rogers arm score was the criterion. Standing height was the only variable found in all five criteria.

6. Shoulder inward rotation strength appeared in three of the multiple correlations; shoulder adduction strength and length of upper arm, in two; and elbow flexion strength in one. Shoulder extension strength was the only variable not found in any of the multiples.

7. In general, the strength variables fared rather poorly in this research. All but one appeared in at least one multiple; however, elbow flexion strength was the only one which ranked first in such a battery (Rogers arm strength).

Discussion

Perhaps the most interesting finding in this research was the high relationship of girth of flexed-tensed upper arm, an anthropometric measure, to strength test variables and arm strength criteria. In interpreting this result, the nature of the subjects should be kept in mind. As described above, they were dominantly mesomorphic, with some tendency toward meso-ectomorphy; and had better than average musculatures. Thus, with few exceptions, the amount of superfluous flesh was slight for most of them; the tape was, therefore, measuring bone and muscle, with a minimum of fat. It is questionable if similar results would be obtained with a random sample of college students.

Another anthropometric test, body height, was also prominent in the multiple correlations. Consequently, it may be concluded that the height of the individual adds significantly to his ability to score well on arm strength measures, as presently conceived.

An interesting aspect of the study is the great difference in multiple correlation obtained with the McCloy and with the Rogers arm strength measures. Contrasting predictive indices, the McCloy formula has 5.88 times more predictive value from these variables than the Rogers arm score. The correlations between the McCloy and Rogers scores, utilizing identical pull-ups and push-ups, is .42.

In examining the scattergram, it was found that heavy subjects with some tendency toward obesity scored high by the McCloy formula and low by the Rogers score; the reverse was true with small, slender subjects. This phenomenon needs further investigation, both to discover the significance of differences that exist between the two arm strength scoring methods and to determine their relative value in the measurement of physical fitness, motor fitness, and general athletic ability.

REFERENCES

1. Clarke, H. Harrison. *A Manual: Cable-Tension Strength Tests.* Chicopee, Mass.: Brown-Murphy Co., 1953.
2. ———. *Application of Measurement to Health and Physical Education,* 2nd Ed. New York: Prentice-Hall, Inc., 1950, chapter 8.
3. ———, Single-Bout Elbow Flexion and Shoulder Flexion Ergography Under Conditions of Exhaustion Testing, *Archives of Physical Medicine and Rehabilitation* **34**: 240 (April 1953).
4. Garrett, Henry E., *Statistics in Education and Psychology,* 4th Edition. New York: Longmans, Green and Co., 1953.
5. MacCurdy, Howard L., *A Test for Measuring the Physical Capacity of Secondary School Boys.* Yonkers, N. Y.: The Author, 1933.
6. McCloy, C. H., A New Method of Scoring Chinning and Dipping, *Research Quarterly* **2**: 132 (Dec. 1931).
7. Seaver, Jay W. *Anthropometry and Physical Examinations.* New Haven: O. A. Dorman company, 1896.
8. Sheldon, W. H., S. S. Stevens, and W. B. Tucker, *The Varieties of Human Physique.* New York: Harper and Brothers, 1940.
9. Mathews, Donald K., Uupublished studies. Springfield College.

(Submitted 7-15-53)

Nation-Wide Survey of Core Curricular Content in Physical Education for College Women

DOROTHY DAVIS
Woman's College of the University of North Carolina
Greensboro, North Carolina

Abstract

In this study 299 schools were surveyed to determine the number having a core curriculum and the various types of courses which were included. From the 198 respondents it was found that 75 schools had no core requirement while 123 have some type of core requirement. The study shows a marked trend toward an area, basic orientation course or specific activities requirement of all participants in the service program. In addition to a basic course, the requirements emphasize body mechanics, team sports, individual sports, swimming, and some form of dance.

THIS REPORT was made to the National Association of Physical Education for College Women by the Subcommittee on Curricular Content in the College Service Program.[1] The study reported here was conducted on a nation-wide scale in an attempt to determine the number of schools having a core curriculum and the various types of courses included.

Procedure

Questionnaires were sent out by five District Chairmen to 299 selected colleges and universities. Each chairman was asked to make as extensive a survey as possible of her district or to use information for her district which was recent and already available. The distribution of the 299 questionnaires according to districts was as follows: Central District, 60; Eastern District, 27; Midwestern District, 48; Southern District, 94; Western District, 70.

Replies to the questionnaire were received from 198 schools (66%). The questionnaire asked if the school had a core requirement in the department. Core requirement was defined as a basic course or courses fulfilling common needs of women students and required of all women non-physical education major students in the school. A respondent replying in the affirmative was asked to describe the contents of the course or courses in outline form.

As shown by Table 1, the number of respondents stating they had no core requirement was 75 schools (38%) while the number with some type of core requirement was 123 schools (62%). It was found that the schools with a core requirement had three general types—area, basic orientation course and specific activity course requirement. Table 1 also shows that 16

---

[1]Other members of the Subcommittee on Curricular Content in the College Service Program are: Ruth Campbell, University of Wyoming, Laramie, Wyoming; Catherine Foland, Hood College, Frederick, Maryland; Grace Fox, Florida State University, Tallahassee, Florida; Sara Houston, Oberlin College, Oberlin, Ohio; and Janet Woodruff, University of Oregon, Eugene, Oregon.

schools had two types of requirement. The results of the answers which were received are presented in the following tables.

Discussion of Results

In Table 2 it is shown that of those 38 schools having an area requirement a larger percentage of them required team sports, an individual or an individual *or* dual sport, swimming (one school permits synchronized swimming to be elected to fulfill the group activity area requirement) and rhythms or some general type of dance. Some requirements allow a choice within several areas, as shown in column—Number with Choice—while others designate definite areas as shown in column—Number with No Choice.

TABLE 1

Specific Requirements Within the Overall Requirement in Physical Education for College Women

N = 198

Type of Requirement	Number	Percentage
No. core requirement	75	38
Area requirement	38	19
Basic orientation course requirement	39	20
Other specific course requirement	62	31
Two types of requirement*	16	8

*NOTE: A total of 214 schools is shown, owing to the fact that 16 schools have two types of requirements.

TABLE 2

Analysis of Definite Area Requirements in the Service Program of Departments of Physical Education

N = 38

Areas	Number with No Choice	Number with Choice	Total	Percentage
Body mechanics	1	1	2	5
Club activities	1	0	1	2.5
Dance—general	7	2	9	23
Folk dance	0	3	3	8
Fundamentals of movement	1	0	1	2.5
Group activities	1	0	1	2.5
Individual sports	15	1	16	42
Individual or dual sports	16	0	16	42
Modern dance	3	5	8	21
Rhythms	16	0	16	42
Sports—general	1	0	1	2.5
Square dance	0	3	3	8
Swimming	19	4	23	60
Swimming test	7	2	9	23
Team sports	32	3	35	92

Table 3 shows the various combinations of area requirements as indicated by the various schools. As an example, three different schools combine an activity from individual or dual sports, team sports, and rhythms for their requirement. The largest percentage require the combination of an activity from individual or dual sports, team sports, rhythms and swimming.

TABLE 3
Analysis of Combinations of Definite Area Requirements in the Service Program of Departments of Physical Education
N = 38

Combinations of Areas	Number	Percentage
Individual or dual sports, team sports, and rhythms	3	8
Individual or dual sports, team sports, rhythms, and swimming	8	21
Body mechanics, team sports, swimming, and choice of dance courses	1	2.6
Body mechanics, individual or dual sports, rhythms and team sports	1	2.6
Club activities, choice of folk, modern or square dance, individual sport, and swimming test	1	2.6
Individual sport and team sport	1	2.6
Two individual sports, modern dance, and team sports	1	2.6
Group activities (synchronized swimming) and individual sport	1	2.6
Individual or dual sports and team sports	1	2.6
Dance-general, individual or dual sports, swimming test, and team sports	2	5
Two individual sports, swimming test, and team sports	1	2.6
Dance-general and sports-general	1	2.6
Dance-general, individual sport, swimming, swimming test, and team sports	1	2.6
Choice of dance-general, swimming, and team sports	1	2.6
Individual sports, rhythms, swimming, and team sports	1	2.6
Individual or dual sport, two rhythms, swimming test, and two team sports	1	2.6
Team sport	1	2.6
Individual sport, modern dance, and team sports	1	2.6
Two individual or dual sports, rhythms, choice of swimming or swimming test, two team sports	1	2.6
Two dance-general, two individual sports, swimming, and team sport	1	2.6
Choice of dance-general, individual sports, swimming, or team sports	1	2.6
Individual sport, modern dance, swimming, and team sports	1	2.6
Dance-general, individual sport, and team sport	1	2.6
Individual sports, choice of modern dance or square dance, swimming, swimming test, and team sport	1	2.6
Rhythms, swimming, two team sports	1	2.6
Choice of modern dance, swimming, swimming test, and team sports	1	2.6
Dance-general, choice of folk, modern or square dance, individual sports, swimming, and team sports	1	2.6
Two fundamentals of movement, swimming, and team sports	1	2.6

As shown by Table 4, of the 39 schools with a basic orientation course requirement, a larger percentage of them include body mechanics, relaxation, posture, daily living activities, departmental policies and procedures, objectives of physical education, weight control, health practices, conditioning exercises and fundamental movement (rhythmic) in their basic course.

As shown by Table 5, of those 62 schools having required specific activity courses, a larger percentage of them required body mechanics, volleyball, basketball, swimming or swimming test, softball, rhythmic fundamentals, folk dance, square dance, stunts and tumbling, health, badminton, and soccer.

TABLE 4
Analysis of Materials Included in Basic Orientation Courses Required in Physical Education Departments
N = 39

Materials Included in Courses	Number	Percentage
Analysis of individual background	2	5
Body mechanics	33	84
Conditioning exercises	10	25
Daily living activities	18	46
Departmental offerings	6	15
Departmental policies and procedures	17	43
First aid	1	2.5
Fundamental movement (rhythmics)	9	23
Fundamental skills	7	18
General orientation	6	15
Guidance in selection of activities	6	15
Gymnastics	3	7.6
Health Practices	11	28.2
Leisure	1	2.5
Modern dance	3	7.6
Motor ability tests	5	13
Objectives of physical education	17	43
Outing activities	1	2.5
Philosophy of physical education	6	15
Philosophy of recreation	5	13
Physical fitness	1	2.5
Posture	22	56
Recreational games	3	7.6
Relaxation	29	74
Rhythmics	6	15
Swimming test	5	13
Swimming	1	2.5
Team sport	4	10
Weight control	15	38

In Table 6 are listed the schools which have two types of requirements—four have an area and basic orientation course combination and 12 have an area and definite course requirement.

As shown by Table 7, of those seven schools which have a requirement but do not consider it to be a core curriculum, a larger percentage require swimming, fundamentals of movement (rhythms), and dance in general.

Table 8 shows that of those schools which do not now have a requirement but have listed the requirements which they feel are important, a larger percentage suggested individual sports, team sports, swimming, body mechanics, and a rhythmic activity.

A more detailed account of some of the basic orientation courses will be submitted at a later date so that any school which wishes to offer such a course may have some suggestions to follow.

Conclusions

The committee feels that the study shows there is a marked trend in the direction of some specific basic curricular material required of all participants

TABLE 5
Aanlysis of Activity Courses Stated as Definite Requirements of Physical Education
N = 62

Specific Activity Courses Required	Number with No Choice	Number with Choice	Total	Percentage
Archery	10	0	10	16
Badminton	13	1	14	22.4
Ballet	1	0	1	1.6
Basketball	25	0	25	40
Body mechanics	34	0	34	54.4
Bowling	7	0	7	11.2
Conditioning exercises	9	1	10	16
Fencing	1	0	1	1.6
First aid	2	0	2	3.2
Folk dance	15	2	17	27.2
Fundamental skills	5	0	5	8
Golf	6	0	6	9.6
Group games	1	0	1	1.6
Gymnastics	5	1	6	9.6
Health	14	0	14	22.4
Hockey	9	1	10	16
Individual sports (no definite sport specified)	5	0	5	8
Lead-up games	4	0	4	6.4
Marching	1	0	1	1.6
Modern dance	10	2	12	19.2
Pyramids	3	0	3	4.8
Recreational games	9	1	10	16
Relaxation	3	0	3	4.8
Rhythmic fundamentals	16	1	17	27.2
Self-testing activities	1	0	1	1.6
Skating	3	0	3	4.8
Skiing	0	1	1	1.6
Soccer	12	1	13	20.8
Social dance	7	0	7	11.2
Softball	18	3	21	33.6
Speedball	6	1	7	11.2
Square dance	15	1	16	25.6
Stunts and tumbling	16	0	16	25.6
Swimming or swimming test	22	1	23	36.8
Tap dance	0	1	1	1.6
Team sports (no definite sports specified)	4	0	4	6.4
Tennis	10	2	12	19.2
Track and field	4	0	4	6.4
Volleyball	26	1	27	43.2

in the physical education service program. These requirements are arranged in three ways: (a) Area; (b) Basic course; (c) Specific activities.

In addition to an apparent trend toward a basic course, the area and specific activity requirement emphasize body mechanics, team sports, individual sports, swimming and rhythmic activities, or some form of dance.

TABLE 6
Analysis of Physical Education Departments with Two Types of Requirements
N = 16

Type of Requirement	Number	Percentage
Area requirement (Table 2)	16	100
Basic orientation course (Table 4)	4	25
Definite courses required (Table 5)	12	75

TABLE 7
Departments Stating a Requirement but Not Considering it a Core Curriculum
N = 7

Type of Requirement	Number	Percentage
Archery	1	14.2
Calisthenics	1	14.2
Conditioning exercises	1	14.2
Dance	2	28.5
First aid	1	14.2
Fundamentals of movement (rhythmic)	2	28.5
Group activity	1	14.2
Group games	1	14.2
Individual sports (no definite sport specified)	1	14.2
Soccer	1	14.2
Swimming	5	71
Team sports (no definite sport specified)	1	14.2
Tennis	1	14.2
Tumbling	1	14.2
Volleyball	1	14.2

TABLE 8
Requirements Suggested by Schools Not Now Having Core Requirement
N = 24

Type of Requirement	Number	Percentage
Achievement of organic fitness	1	4.2
Body mechanics	9	37.5
Dance	6	25
Dual sports	2	8.3
Departmental offerings	1	4.2
Fundamentals of movement	2	8.3
Group activities	1	4.2
Health Education	3	12.5
Individual sports (no definite sport specified)	15	62.5
Modern dance	1	4.2
Orientation course	3	12.5
Recreational skills	5	20.8
Relaxation	1	4.2
Rhythmic activity or fundamentals	9	37.5
Sampling of skills and activities	2	8.3
Swimming or swimming test	10	42
Team sports (no definite sport specified)	14	58.3

(Submitted 7-11-53)

Prevalence of Certain Harmful Health and Safety Misconceptions Among Fifth- and Sixth-Grade Children

JOSEPH G. DZENOWAGIS
Boston University, Sargent College
Boston, Masachusetts

LESLIE W. IRWIN
Boston University, School of Education
Boston, Masachusetts

Abstract

In this study an instrument was constructed and evaluated for determining the prevalence of certain harmful health and safety misconceptions among fifth and sixth grade children. The instrument was used to determine the prevalence of 216 harmful health and safety misconceptions among 2,210 fifth grade children and 1,881 sixth grade children in six urban and suburban communities in Eastern Massachusetts representing both industrial and residential areas. The analysis of the results revealed that many harmful health and safety misconceptions were subscribed to by the fifth and sixth grade children participating in this investigation.

THE PURPOSE OF this study was: (1) to construct and evaluate an instrument for determining the prevalence of certain harmful health and safety misconceptions among fifth- and sixth-grade children; and (2) to determine the prevalence of responses to this instrument which suggest the occurrence of certain harmful health and safety misconceptions among fifth- and sixth-grade children, and variations of this prevalence according to grade level. Misconceptions are erroneous ideas, notions or beliefs. This investigation is concerned with two kinds of misconceptions: health misconceptions and safety misconceptions. For the purpose of this study they are expressed in declarative sentences. Health misconceptions are expressed as statements pertaining to health that lack a scientific basis and are not in accord with current authoritative medical thought. Safety misconceptions are expressed as statements pertaining to safety that lack a scientific basis and are not in accord with current authoritative thought in the field of safety.

Justification of the Study

Education has the challenging task of helping individuals to acquire concepts of healthful and safe living which are essential to, or consistent with, desirable behavior. It also has the task of helping individuals to identify and correct their acquired misconceptions pertaining to health and safety which are not consistent with such behavior. The accomplishment of these two tasks depends to a great extent upon the intelligent selection and organization of subject matter and experiences by the persons responsible for the education in our democracy.

Merrill (13), Staton (19), and Boyd (3) have provided important information to educators concerning those concepts of healthful living which are of functional value in contributing to the general education of the individual. This information is of great assistance in the planning of health instruction programs. More effective planning, however, would be possible if significant data were made available to educators concerning the prevalence of harmful health and safety misconceptions among school children and variations of this prevalence according to grade level.

Review of Related Literature

A review of the research revealed a few studies (15, 16, 17, 18) related in purpose and method to the one proposed here. However, the review of the indexed literature did not reveal any studies that were concerned with harmful health and safety misconceptions of fifth- and sixth-grade children.

Numerous books (1, 4, 5, 8, 11, 20) and articles (2, 6, 7, 9, 23) dealing with health superstitions, health fallacies, nostrums, health fads, and quackery were discovered in the literature. They served as valuable sources for ideas in the initial steps of compiling statements of misconceptions for the instrument used in this investigation. These books and articles are primarily concerned with the enumeration and explanation of unfounded health beliefs and may well serve as references for teachers to help in the elimination of health and safety misconceptions.

The review of related research revealed only a few studies (15, 16, 17, 18) that endeavored to investigate the prevalence of health misconceptions and treat the data statistically. These investigations indicated that prospective teachers, secondary school students, and young adults as a whole subscribe to health misconceptions to a surprising degree.

Face validity, or verification of certain statements as health misconceptions in these investigations, was obtained by submitting them to medical and subject-matter experts for evaluation or by reference to authoritative sources.

Nearly all the investigators used true-false questionnaires for the determination of prevalent health misconceptions. Rhoton (15) and Salt (17), using the Spearman-Brown Prophecy formula, obtained reliability coefficients of .817 and .903 respectively for the true-false tests used in their investigations.

Hancock (10) used a jury of experts to evaluate certain popular science misconceptions according to their potentialities for affecting the behavior of individuals subscribing to them and concluded that the misconceptions having the most influence on behavior were those related to health.

The analysis of related literature strongly indicates the urgent need for reliable and valid instruments that can be used to obtain data concerning the prevalence of harmful health and safety misconceptions in our school population. This investigation was an attempt to meet this need by constructing and evaluating one such instrument for use among fifth and sixth grade school children.

Procedure

Criteria for the Development of the Instrument. The securing of data for this study depended upon the development of an instrument for the determination of certain harmful health and safety misconceptions and its administration to the children comprising the entire population of the fifth and sixth grades of six communities. It was assumed that such an instrument, in order to be satisfactory, should conform to the following criteria:

1. It should be a valid and reliable measure of the prevalence of certain harmful health and safety misconceptions.
2. It should be easy and quick to administer.
3. It should provide for the expression of different degrees of credulity and an opportunity for the examinee to indicate those items which are not known or understood.
4. It should be free from statements pertaining to religion or sex.
5. It should be suited to the examinees' reading level.
6. It should give no outward indication that a study of health and safety misconceptions is being conducted.
7. It should allow for the presentation of a maximum number of health and safety misconceptions.
8. It should lend itself to machine scoring.
9. It should be short enough to prevent the examinees from becoming fatigued and disinterested.

Compilation of Health and Safety Misconceptions. In order to develop items for the instrument, a list of 203 health misconceptions and 30 safety misconceptions was compiled that might be considered harmful to an individual or individuals if used as guides for behavior. The sources utilized for this compilation included: (1) previous studies of health misconceptions; (2) published books and articles pertaining to health and safety misconceptions; (3) contributions from various students, teachers, and physicians; and (4) press, magazine, radio, and other forms of advertising which stated or implied health misconceptions.

Formulation of the Items in the Preliminary Instrument. The next step in the development of the instrument was that of writing a suitable statement of each misconception to be included in the instrument. To facilitate this process, the following criteria were used as guides in the formulation of each statement:

1. It should be a declarative sentence.
2. It should be brief.
3. It should be completely false.
4. It should have only one interpretation.
5. It should be suited to the reading level of fifth and sixth grade children.

During the formulation of these statements, several fifth-grade children were interviewed to obtain information about the complexity, clarity, and difficulty of each item. Analysis, criticism, and suggestions for revision of the statements were also obtained from members of the Measurement and Evaluation Seminar during the 1951 Summer Session at the School of Education, Boston University. In addition, the statements were submitted to four elementary school specialists during personal interviews to obtain further

analysis, criticism, and suggestions for revision of the statements in accordance with the criteria set for their formulation.

Jury Validation and Evaluation of the Health and Safety Misconceptions. The validation and evaluation of the health and safety misconceptions that were used in this investigation required the development of a special evaluation instrument and its administration to medical specialists and subject-matter experts in the areas of health and safety.

Experts Co-operating in the Evaluation of Health Misconceptions. Julius W. Fryer, M. D., Ph. D., clinical director of the Danvers State Hospital, Hathorne, Massachusetts; Roger Osterheld, M. D., superintendent of the Monson State Hospital, Monson, Massachusetts; Robert W. Keeler, M. D., Attleboro, Massachusetts; Otto S. Nau, Jr., M. D., Arlington, Massachusetts; Carl J. De Prizio, M. D., Attleboro, Massachusetts; Ronald W. Rutherford, M. D., Newton Centre, Massachusetts; William Stobbs, M. D., Attleboro, Massachusetts; John Mulligan, M. D., Bridgewater, Massachusetts; Bradford Lawrence, M. D., Attleboro, Massachusetts; Max E. Rukes, M. D., Charlestown, Massachusetts; Claire E. Turner, D. Sc., Dr. P. H., assistant to the President, National Foundation of Infantile Paralysis, New York, New York; Wesley Staton, Ed. D., associate professor of health and physical education, University of Florida, Gainesville, Florida.

Experts Co-operating in the Evaluation of Safety Misconceptions. Herbert J. Stack, Ph. D., director, Center for Safety Education, New York University, New York, New York; Wayne P. Hughes, director, School and College Division, National Safety Council, Chicago, Illinois; Vivian Weedon, Ph. D., staff member, School and College Division, National Safety Council, Chicago, Illinois; Marion Telford, staff member, School and College Division, National Safety Council, Chicago, Illinois; Russell Brown, staff member, School and College Division, National Safety Council, Chicago Illinois; Renato Leonelli, Ed. D., associate professor of education, Rhode Island College of Education, Providence, Rhode Island.

Two psychiatrists, two pediatricians, one surgeon, five general practicing physicians, and three health educators co-operated in the validation and evaluation of the health misconceptions. The safety misconceptions were validated and evaluated by seven subject-matter specialists in safety education. Each of the experts received instructions to:

1. Rate only the misconceptions that were completely false by checking one of the following: (1) not harmful, (2) slightly harmful, (3) moderately harmful, (4) very harmful, or (5) extremely harmful.
2. Delete any statement that was not completely false.
3. Rate each statement only according to the degree of harm which such a misconception would engender, and not according to prevalence or possibility of occurrence.

In the analysis and treatment of the data that was obtained from the jury validation and evaluation of health and safety misconceptions it was necessary to classify the completely false health and safety misconceptions on the basis of their median ratings as being slightly harmful, moderately harmful,

very harmful or extremely harmful. The 17 statements which were deleted by one or more of the jury members as being not completely false were not used in the remainder of the investigation.

Construction and Administration of the Preliminary Instrument. During the spring of 1952, a preliminary instrument was constructed and administered to 152 children comprising the entire population of the fifth and sixth grades in one community. The preliminary instrument consisted of four inventory forms. Each form contained ninety statements. Approximately two-thirds of the statements were false and one-third of the statements were true in each of the inventory forms. The true statements were used to serve as camouflage for the misconceptions. They were needed in order to prevent the examinees from becoming aware that a study of health and safety misconceptions was being conducted.

Each examinee was provided with a Bugraphic pencil and an International Business Machine answer sheet with directions and sample exercises overprinted on one side and provisions for responses on the other. The provisions for responses allowed the examinees to respond to each statement in a particular form in one of the following ways: true, sometimes true, false, don't know, or don't understand. The examinees were instructed not to guess and were allowed as much time as they needed to complete the particular form they were using.

The four inventory forms were distributed randomly and equitably among the examinees in each of the fifth and sixth grades. Homeroom teachers were supplied with detailed directions pertaining to the administration of the inventory forms. The forms were administered by homeroom teachers in regular classroom situations. The information obtained during the administration of the preliminary forms revealed some of the revisions needed for the improvement of the instrument and its administration.

Construction of the Final Instrument. One hundred eighty-seven health misconceptions that were considered to be harmful and completely false by the juries of experts were distributed in approximately equal numbers among five inventory forms. True statements or camouflage items were manually distributed among the misconceptions. The same 17 camouflage items were used in each form with the exception of Form E which contained only 16 camouflage items. Forms A, B, C, and D contained 43 misconceptions. Form E contained 44 misconceptions. Thus each inventory contained 60 statements. The answer sheets and directions to teachers used with the final inventory forms were similar to the ones used in the preliminary instrument except for a few minor changes in typographical format.

Administration of the Final Instrument. The final instrument was administered to 2,210 fifth-grade children and 1,881 sixth-grade children comprising the total fifth- and sixth-grade population in six urban and suburban communities in eastern Massachusetts representing both industrial and residential areas. The forms were equally and randomly distributed in the total

population at each grade level by classroom units and were administered under the direction of homeroom teachers during regular class hours.

Organization and Treatment of Data. The data obtained as a result of the administration of the final instrument were analyzed to indicate:

1. The index of discrimination of each misconception at the sixth-grade level.
2. The median, mean, standard deviation, and reliability of each inventory form at each grade level.
3. The percentages of all responses to each misconception by fifth- and sixth-grade children.
4. The variation of the "true" and "sometimes true" responses to each misconception by fifth and sixth grade children.
5. The variation of the "don't know" and "don't understand" responses to each misconception by fifth- and sixth-grade children.

It was assumed that the sampling was random and that comparable groups responded to the statements in the inventory forms.

Evaluation of the Instrument

The instrument used in this investigation may be evaluated by answering the following questions:

1. Was the instrument easy and quick to administer?

Yes, a number of conferences with various teachers who administered the instrument revealed that the average amount of time for the pupils to read the directions, to respond to the sample items, and to complete an inventory form was approximately 25 minutes. There were no complaints concerning the administration of the instrument from any of the teachers and principals who co-operated in this investigation.

2. Did the instrument function as intended in the determination of the prevalence of harmful health and safety misconceptions among fifth and sixth grade children?

Yes, the instrument functioned as intended. There were two indications that the fifth- and sixth-grade children were honest in their responses to the statements in the instrument. First, the frequency or error for the true statements was small, approaching zero, since they were functioning as camouflage items. Second, the "don't know" and the "don't understand" responses to the statements containing difficult terminology were high in frequency. Therefore, it may be assumed that the fifth- and sixth-grade children were not aware that a study of misconceptions was being conducted. In addition, it is reasonable to assume that there was a minimum of guessing by the fifth- and sixth-grade children responding to statements in the instrument.

3. Has the instrument substantial validity?

Yes, the validity of the instrument was established in the following ways:

(a) The 187 health misconceptions used in the final instrument were considered to be completely false by all members of the jury of experts that co-operated in the health misconception evaluation.

(b) The 29 safety misconceptions used in the final instrument were considered to be completely false by all members of the jury of experts that co-operated in the safety misconception evaluation.

(c) The true statements used in the final instrument were obtained from Merrill's (13) list of concepts of healthful living. Seventeen statements were selected for use in the final instrument on the basis that 90 per cent or more of the examinees in the trial study had responded to them correctly.

(d) Flanagan Indices of Discrimination (21) were determined for each misconception in each inventory form at the sixth grade level. Only seven of the 216 misconceptions used in the final instrument had indices below the acceptable minimum of 0.20. Two hundred and one of the 216 misconceptions used in the final instrument had indices of 0.25 or over and may be considered as outstandingly valid items (21). Therefore, it may be assumed that the discriminating power of the instrument as a whole is adequate.

4. Were the five inventory forms used in this investigation reliable instruments for obtaining information about the prevalence of certain harmful health and safety misconceptions among fifth and sixth grade children?

The inventory forms used in this investigation have limited reliability. Reliability coefficients of the five inventory forms ranged from 0.74 to 0.82. They were determined by the use of Kuder-Richardson formula #20 (12).

Conclusions

The following conclusions are based upon data obtained from this investigation.

1. Many harmful health and safety misconceptions are very prevalent among fifth- and sixth-grade children.

(a) Fifty per cent or more of the fifth-grade children subscribed to 72 of 216 harmful health and safety misconceptions.

(b) Fifty per cent or more of the sixth-grade children subscribed to 69 of 216 harmful and safety misconceptions.

2. Variations in the prevalence of certain harmful health and safety misconceptions do exist.

(a) Eighty-five of 216 harmful health and safety misconceptions increased in prevalence from the fifth to the sixth grade. The percentage increase of the prevalence of these 85 misconceptions ranged from 1 per cent to 16 per cent.

(b) One hundred twenty-two of 216 harmful health and safety misconceptions showed a decrease in prevalence from the fifth to the sixth grade. The percentage decrease of the prevalence of these 122 harmful health and safety misconceptions ranged from 1 per cent to 23 per cent.

(c) Nine harmful health and safety misconceptions showed no change in prevalence between the fifth and sixth grade.

3. The instrument used in this investigation has limited reliability. However, it is easy and quick to administer, has substantial validity, and functions as intended for the determination of the prevalence of 216 harmful health and safety misconceptions among fifth- and six-grade children.

Recommendations

On the basis of the results of this investigation the following recommendations seem justifiable:

1. Fifth- and sixth-grade teachers should identify the harmful health and safety misconceptions subscribed to by their pupils as a partial basis for the selection and organization of subject matter and experiences in the area of health and safety.
2. Health and safety instruction in the fifth- and sixth-grades should be partially graded from year to year on the basis of the variations of the prevalence of harmful health and safety misconceptions.
3. The misconceptions used in this investigation should be revaluated periodically by juries of experts.
4. More extensive studies of health and safety misconceptions involving other grade levels should be made.

APPENDIX

TABLE 1

Variations of the "True" and "Sometimes True" Responses of Fifth- and Sixth-Grade Children to Certain Extremely Harmful Health and Safety Misconceptions

Extremely Harmful Health and Safety Misconceptions	Percentage Frequency of Responses	
	Grade 5	Grade 6
	T and ST	T and ST
(1)	(2)	(3)
1. The best doctors always promise to make people healthy	73	67
2. The only good way to help a drowning person is to jump in the water to save him	72	68
3. It is usually safe to go in swimming alone if you know how to swim	60	45
4. Oil, grease, and gas fires should be put out with plenty of water	57	48
5. A bullet cannot go off unless it is fired by a gun	54	43
6. If your clothing catches fire, you should always run for water	53	43
7. Bicycle riders should ride on the left hand side of the road to be safe	50	52
8. Most mental sicknesses cannot be helped by any treatment	47	51
9. It is a good idea to make an unconscious person drink something	42	50
10. A person having a stomach ache should usually take a laxative	41	42
11. It is best to go to doctors who advertise in newspapers	40	31
12. It is impossible to cure any cancer	40	36
13. It is all right to point a gun at someone if you are sure that it is not loaded.	39	29
14. A person always comes up to the top of the water three times before he drowns	38	41
15. All mad dogs foam at the mouth	37	37
16. Tuberculosis is a shameful disease to have	35	30
17. A good way to treat a burn is to put iodine on it	30	25
18. It is safe to cross the street without looking when the traffic light is yellow and red	29	30
19. Dynamite caps are always safe unless fastened to a fuse	26	24
20. It is safe to cross the street without looking when the traffic light is red.	25	26
21. A person who has recovered after having tuberculosis cannot get it again.	25	20
22. Throwing oil or gasoline on a slow fire is a wise thing to do	23	14
23. You should be ashamed if anyone in your family is mentally ill.	20	13
24. Bicycle riders do not have to obey traffic lights	20	15

Extremely Harmful Health and Safety Misconceptions	Percentage Frequency of Responses	
	Grade 5	Grade 6
	T and ST	T and ST
(1)	(2)	(3)
25. Unless someone in your family has tuberculosis there is no chance you will get it	19	18
26. All laxatives are safe to use regularly	19	22
27. Touching a light switch or light chain with wet hands is not dangerous.	18	14
28. When you are riding a bicycle you never have to use hand signals...	18	12
29. It is all right to use sleeping pills without a doctor's advice	17	15
30. If you meet a dog that frightens you, it is always best to start running.	16	8
31. It is safe to cross the street without looking when the traffic light is green.	15	9
32. People have accidents only when their "number" is up	13	8
33. Throwing a person into deep water is a good way to teach him to swim.	12	11
34. The eyes can be made stronger by looking at the sun	12	9
35. When you are swimming it is a good joke to call for help when you don't need it	9	3

TABLE 2

Variations of the "True" and "Sometimes True" Responses of Fifth- and Sixth-Grade Children to Certain Very Harmful Health and Safety Misconceptions

Very Harmful Health and Safety Misconceptions	Percentage Frequency of Responses	
	Grade 5	Grade 6
	T and ST	T and ST
(1)	(2)	(3)
1. The best way to get a tan is by sleeping in the sun	80	66
2. There are some pills that people can take which will cure the common cold.	79	78
3. It is always impossible for a person with cramps to swim	75	76
4. There are certain cough medicines that will cure and prevent the common cold	72	79
5. Spring water that is clear and cold is always safe for drinking	72	67
6. Most fat people are very healthy	68	69
7. If you have any disease or sickness you will always feel some pain...	67	65
8. All people with rosy complexions are very healthy	66	68
9. Every disease needs a drug or medicine for its cure	65	64
10. Any food that does not smell or taste spoiled is safe to eat	60	66
11. Iodine is the best treatment for wounds caused by stepping on rusty nails.	59	67
12. All radio advertising about what is good or bad for your health is true.	59	61
13. People should walk on the right hand side of the road if there are no sidewalks	57	55
14. It is always safe to drink water which has just been taken from a well or spring	57	47
15. A good safety rule for bicycle riders is: "Ride on the sidewalk as much as possible."	55	34
16. All health advertisements in papers and magazines are true	54	57
17. All advertising on television about what is good or bad for health is true.	52	56
18. Persons can clean their blood by eating certain foods	49	50
19. A great deal of exercise can never hurt anyone	48	55
20. Most accidents cannot be prevented	48	42

Very Harmful Health and Safety Misconceptions	Percentage Frequency of Responses	
	Grade 5 T and ST	Grade 6 T and ST
(1)	(2)	(3)
21. Mental illness usually happens suddenly	48	43
22. People who exercise a lot live longer than other people	47	57
23. There are special laxatives that will help prevent or cure the common cold.	41	43
24. All children with heart murmurs will surely have heart trouble later on in life	40	32
25. An exercise is not good unless it makes your muscles sore and stiff	37	14
26. Blowing your nose as hard as you can is not harmful	36	24
27. Most insane persons were born insane	36	38
28. Most people who get tuberculosis will die in a short time	34	43
29. Baby teeth need very little care because they will soon fall out	32	28
30. Houseflies are harmless because they are unable to bite	31	28
31. The first thing to do in treating a burn is to put cold water on it	31	22
32. Eating meat more than once a day is harmful to most persons	31	34
33. The best way to get water out of your ears after swimming is to hold your nose and mouth closed and blow hard	30	22
34. Drinking raw milk fresh from the cow is a very healthy thing to do	29	30
35. All persons should take laxatives whenever they are constipated	25	30
36. A good way to take care of blisters is to pinch a hole in them	25	25
37. If you feel all right, you can be sure that you do not have tuberculosis	25	18
38. You should never eat when you are sick because you feed the disease	24	21
39. Everyone should take a laxative once a week	23	27
40. Fresh raw milk is a better food for your health than pasteurized milk	23	25
41. Measles is never harmful	22	26
42. Sickness is usually punishment for being bad	22	19
43. It is possible to tell what is going to happen to people from their dreams	20	36
44. You don't need to worry about having tuberculosis unless you are coughing a lot	19	18
45. Some people should drink very little water because it turns to fat in their bodies	19	14
46. It is safe to use toothpicks or matchsticks for removing wax from ears	18	14
47. When tuberculosis is getting started a person always has a pain in the chest	17	22
48. Squeezing the pus out of boils and pimples with your fingers is good for your health	17	16
49. Whooping cough is never harmful	17	18
50. Smart children usually die at an early age	17	12
51. Teeth need care only when they ache	15	9
52. The best way to remove pus from boils and pimples is with your fingers	15	14
53. Looking into the sun can never hurt your eyes	14	12
54. A good way to treat frostbite is to rub the frostbitten part with snow	14	14
55. Most illnesses are caused by constipation	14	18
56. All persons would be healthier if they ate only raw food	14	21
57. Fat people can feel quite sure that they will never get tuberculosis	12	11
58. Eating little or no breakfast is a good health habit for all people	10	9
59. The best place for shelter, during a thunderstorm, is under a tree	7	6

TABLE 3

Variations of the "True" and "Sometimes True" Responses of Fifth- and Sixth-Grade Children to Certain Moderately Harmful Health and Safety Misconceptions

Moderately Harmful Health and Safety Misconceptions	Percentage Frequency of Responses: Grade 5 T and ST	Percentage Frequency of Responses: Grade 6 T and ST
(1)	(2)	(3)
1. Brushing your teeth every day is a sure way of stopping decay	95	97
2. Most fat people are happy and jolly	84	80
3. Nose drops will cure a cold which is causing a stuffy nose	81	81
4. Any person who feels all right is sure to be in good health	78	68
5. Taking vitamin pills will guarantee you good health	72	76
6. Wearing bathing hats or ear plugs while swimming will give a person complete protection for his ears	71	74
7. All persons should use nose drops and mouth washes daily when they have a cold	69	65
8. There are some pills that people can take which will prevent the common cold	68	75
9. Wearing sunglasses will give your eyes complete protection from the sun	68	71
10. Most persons who look thin are certain to be underweight and in poor health.	68	68
11. Taking vitmain pills is the best way to get your necessary vitamins	67	64
12. The use of skin lotions is a healthful way to make any skin beautiful	67	65
13. It is necessary to go to a doctor only when you feel sick	67	52
14. Wearing eyeglasses will always make a person's eyes stronger	65	68
15. Everyone who has weak feet should wear arch supports to strengthen them	64	70
16. If your eyes do not hurt, you can be sure they are healthy	64	57
17. You can be sure anything a scientist says about health is true	63	60
18. Bad breath can be stopped for good by using special mouth washes	62	63
19. Most colds can be cured by taking vitamin pills	62	57
20. Pain near the heart is generally a sign of heart disease	60	62
21. Any food that smells and tastes good is safe to eat	60	51
22. Sugar diabetes is caused by eating too much sugar	59	64
23. A daily bowel movement is always necessary so a person can stay healthy	59	68
24. The use of tooth powders and pastes is sure to make a person's gums firm	59	69
25. The use of skin creams and lotions will make any skin clear and healthy	59	58
26. The only good way to lose weight is by exercising	58	61
27. Cotton should be the first thing put on a cut to stop the bleeding	57	45
28. Mouth washes are sure to prevent or cure diseases of the mouth and throat	56	59
29. The only good treatment for weak arches is to have arch supports placed in the shoes	55	65
30. Expensive food is always the best food to eat	54	52
31. Eating between meals causes most children to have poor health	53	56
32. The vitamins in certain pills are better than the vitamins in natural foods	53	57
33. Unlucky people are sure to fail at the new things that they try to do	53	40
34. Any exercise is bad for persons who have heart trouble	52	61
35. You can always tell if a dog is friendly by his looks	52	40
36. The best medicines are the medicines that taste the worst	50	56
37. A fortune teller can tell your future by looking at the lines in the palm of your hand	48	40
38. Skipping one or two meals a day is a healthy way to get thin	44	24
39. Some persons have the ability to tell your fortune	44	40
40. A good way to help a person get rid of the hiccoughs is to frighten him	44	53
41. Most persons need big muscles in order to be healthy	44	38

Moderately Harmful Health and Safety Misconceptions	Percentage Frequency of Responses	
	Grade 5 T and ST	Grade 6 T and ST
(1)	(2)	(3)
42. Good doctors usually advertise	43	41
43. Smoking is not harmful because many doctors and athletes smoke	43	40
44. Most colds cannot be prevented	41	42
45. An all vegetable diet is the natural and best diet	41	37
46. Perons who have pimples or boils usually have bad blood	40	46
47. Sunburns are harmless even when they are painful	39	33
48. Most cases of baldness can be cured if treated early	38	40
49. Good health does not depend on what you eat	38	41
50. Food that tastes good is usually bad for your health	38	42
51. Nighttime is the only time that one ever needs sleep or rest	37	32
52. It is always good for your health to eat overripe fruits	37	31
53. Exercising regularly is a sure way to prevent disease	36	38
54. Cancer is catching	35	34
55. If a person wants to be strong and healthy, he should eat plenty of raw meat	34	24
56. The best way to treat a black eye is to put a piece of raw meat on it	33	40
57. Drinking water with your meals is always bad for your health	33	31
58. Bananas should be kept out of a good diet because they are hard to digest.	32	33
59. The best way to brush your teeth is sideways	32	23
60. Only bad smelling odors can be harmful to your health	31	26
61. All cosmetics are healthful to use	30	30
62. All sick people should drink bottled mineral water to bring back their health	29	26
63. Missing a bowel movement for one day is always a sign of constipation	28	33
64. Most people who have tuberculosis were born with it	28	29
65. People should use headache pills every time they have a cold	28	29
66. Cheese is a bad food to eat because it is hard to digest	27	28
67. The first and best thing to do in caring for a cold is to take a laxative	27	27
68. People can never change their food likes and dislikes	26	28
69. Wanting to eat candy and sweets is always a sign that your body needs sugar	26	24
70. A person's health depends mostly on his luck	26	10
71. Cheese should be kept out of a good diet because it is constipating	25	18
72. Eating fruits and vegetables at the same meal is a bad health practice	25	24
73. All vegetables and fruits should be eaten raw	25	22
74. It is very hard for thin persons to keep from getting tuberculosis	24	29
75. If you are hungry most of the time, you can be sure you have a tapeworm	24	16
76. A cold can usually be cured by eating raw onions	24	25
77. All medicines that have alcohol in them are harmful	23	22
78. Adding certain bath powders to the bath is a healthful way to lose weight	22	17
79. All children with heart murmurs are sickly	21	22
80. It is generally a good idea to have a radio in your bathroom	21	28
81. A good health rule for all people to follow is: "Eat only the foods you like best"	20	14
82. Wearing eyeglasses will always make a person's eyes weaker	20	19
83. A good health rule to follow is: "Feed a cold and starve a fever"	19	21
84. If you break a mirror you will have seven years of bad luck	18	15
85. When you walk or run it is best to point your toes out toward the side	15	11
86. A good way to treat a black eye is to press the eye with a knife handle	12	11
87. A good way to help digest your food is to smoke a cigarette after you eat	12	10
88. A person's future is determined by the star under which he is born	12	10

TABLE 4

Variations of the "True" and "Sometimes True" Responses of Fifth- and Sixth-Grade Children to Certain Slightly Harmful Health and Safety Misconceptions

Slightly Harmful Health and Safety Misconceptions	Percentage Frequency of Responses	
	Grade 5	Grade 6
	T and ST	T and ST
(1)	(2)	(3)
1. To go on a diet always means to eat less food	80	76
2. Everyone who is on a diet is trying to lose weight	80	74
3. Any person who sees clearly can be sure he doesn't need glasses	80	76
4. The use of tooth powders or pastes will always cure a persons bad breath	74	79
5. Some people are born lucky	71	64
6. Most colds can be prevented by taking vitamin pills	69	69
7. Men with large muscles are always healthier than men with small muscles	68	69
8. Using a toothpick is the best way to get things from between your teeth	67	63
9. A mouth wash is healthful because it helps kill germs in the mouth and throat	65	73
10. There are no living germs in pasteurized milk	58	57
11. Anyone who keeps his skin clean will never have pimples	58	56
12. People are born with their food likes and dislikes	50	51
13. Persons should eat only when they feel hungry	50	42
14. A pain in your right side usually means that you have appendicitis	49	49
15. People should protect themselves from catching cold by gargling with a mouth wash	42	46
16. Fish is a food that is very good for the brain	42	44
17. Cutting or shaving a person's hair makes it grow faster and thicker	41	51
18. Honey is a good food for sweetening a sour stomach	41	37
19. Most dogs do not remember the people who were mean to them	41	34
20. It is a bad health habit to drink water while you exercise	38	49
21. Bananas should be kept out of a good diet because they make people fat	38	39
22. White bread that is enriched with vitamins is a better food than whole wheat bread	38	36
23. Milk is pasteurized to make it easy to digest	36	29
24. Friday the Thirteenth is an unlucky day for most people	34	43
25. Eating two or more different kinds of fruits during the same meal is a bad health practice	32	29
26. Swallowing the seeds of fruits generally causes appendicitis	32	34
27. Persons who open umbrellas indoors will bring themselves bad luck	32	28
28. You will have bad luck if a black cat crosses the path in front of you	32	26
29. Drinking milk while you are eating fish is a bad health practice	27	23
30. A good health rule to follow is: "Do not eat fruits that have been mixed with milk"	25	22
31. Persons can always prevent pimples by eating more raw foods	25	18
32. Sleeping on your left side is bad for your heart	23	19
33. Baldness is usually caused by wearing hats	18	21
34. Some houses are visited by ghosts	7	5

REFERENCES

1. BAUER, W. W., *Health, Hygiene, and Hooey*, Indianapolis: Bobbs-Merrill Co., 1938.
2. BINGER, CARL, Medical Information and Misinformation, *Mental Hygiene*, **31**: 1-13 (Jan. 1947).
3. BOYD, CLIFFORD A., *A Determination of Concepts of Healthful Living and Their Relative Importance for a Course in College Health.* Unpublished Doctor's dissertation, Boston University, 1949.

4. Caldwell, Otis W., and Gerhard E. Lundeen, *Do You Believe It?*, New York: Doubleday Doran and Co., 1934.

5. Fishbein, Morris, *Shattering Health Superstitions*, New York: Liveright Publishing Corp., 1934.

6. Fluck, Paul H., Thirty Million People Can Be Wrong, *Today's Health*, **12**: 42-43 (Dec. 1950).

7. Forsythe, Warren E., Thirty-one Fallacies About Health, *Hygeia*, **4**: 512-513, 544-545 (April 1946).

8. Furcolowe, Charles, *This Will Kill You*, New York: B. C. Forbes and Sons, 1951.

9. Gannon, James A., Medical Misconceptions, *Hygeia*, **27**: 170-171, 205-208 (March 1949).

10. Hancock, Cyril H., *An Evaluation of Certain Popular Science Misconceptions*. Unpublished Master's thesis, Colorado State College of Education, 1938.

11. Kallet, Arthur, and F. J. Schlick, *One Hundred Million Guinea Pigs*, New York: The Vanguard Press, 1933.

12. Kuder, G. F., and M. W. Richardson, The Calculation of Test Reliability Coefficients Based Upon the Method of Rational Equivalence, *Journal of Educational Psychology*, **30**: 681-687 (Dec. 1939).

13. Merrill, Charles D., *A Determination of Concepts of Healthful Living Which are of Functional Value in Contributing to the General Education of Elementary School Pupils*. Unpublished Doctor's dissertation, Boston University, 1949.

14. Peatman, John G., *Descriptive and Sampling Statistics*, New York: Harper and Brothers, 1947.

15. Rhoton, Paul, *Health Misconceptions of Prospective Teachers*. Unpublished Doctor's dissertation, Pennsylvania State College, 1932.

16. Robb, Mary Jane, *A Study of the Relative Prevalence of Unfounded Health Beliefs in a Particular Rural Area*. Unpublished Master's thesis, Ohio Wesleyan University, 1939.

17. Salt, Benton E., *A Study of the Relative Prevalence of Certain Health Misconceptions and Superstitions as Subscribed to by Boys and Girls in Florida Public Schools*. Unpublished Doctor's dissertation, New York University, 1936.

18. Sanchez, Genevieve, *Commonly Held Superstitions Which May Affect Health Practices*. Unpublished Master's thesis, Colorado State College of Education, 1937.

19. Staton, Wesley M., *A Determination of Fundamental Concepts of Healthful Living and Their Relative Importance for General Education at the Secondary-School Level*. Unpublished Doctor's dissertation, Boston University, 1949.

20. Thomen, August A., *Doctors Don't Believe It—Why Should You?*, New York: Simon and Schuster, 1941.

21. Thorndike, Robert L., *Personnel Selection*, New York: Harper and Brothers, 1947. Pp. 347, 245.

22. Walker, George A., and Eleanor Saltzman, False Health Notions. *Hygeia*, **20**: 32-34, 59 (Jan. 1942).

23. Weisinger, Mort, Safety Rules That Can Kill You. *This Week Magazine*, pp. 4-7 (Aug. 6, 1950).

(Submitted 8-8-53)

Time-Velocity Equations and Oxygen Requirements of "All-Out" and "Steady-Pace" Running[1]

FRANKLIN M. HENRY
University of California
Berkeley, California

Abstract

Oxygen requirement is visualized theoretically as the sum of two exponential terms—oxygen intake and debt—both functions of the time required for running a unit distance. Estimates of requirement based on world record times are consistent with this theory and with measured requirements at lower speeds. An equation of motion, derived from acceleration and fatigue factors, describes the position and speed of the runner at any time t in either steady-pace or all-out runs. It has been verified by data on 24 track men and 30 inexperienced athletes who were timed with an automatic recorder on dashes and 300-yd. runs. The metabolic cost of running, computed from these two mathematical models, is considerably less for steady pace than for all-out or certain other coach-recommended velocity patterns. Percentage of time saved *decreases* with increased velocity. Although this finding disagrees with previous concepts, it is explainable on theoretical grounds. Tables and curves are given for simplified computation of time saved. Hand or foot reaction time and sprint speed are uncorrelated ($r = 0.09$ and 0.04).

A. V. HILL, as early as 1927, advocated the avoidance of higher-than-average velocity in all runs longer than 100 yards because of the excessive metabolic cost of high speeds (10). About the same time, Sargent obtained oxygen requirement data on one runner in the speed range 5-9 yd/sec, and derived the well-known 3.8th power law (11). In 1950, Christensen and Högberg recommended that the middle distance runner should maintain his average speed and avoid "rushing" which is apt to bring into play the metabolically expensive lactate debt mechanism (1). Their subject, an experienced middle distance runner, was observed to show a loss of efficiency at the relatively slow velocity of 5 yds/sec. This loss was accelerating rapidly at the highest speed studied, 6 yds/sec, which is still a rather slow speed in track competition (2). More recently, De Moor has given statistical evidence of the relation between large lactate oxygen debts and reduced efficiency (3).

Earlier articles in the present series have considered the equation of motion during the acceleration phase of a sprint run (9), the oxygen requirement of running within the velocity range 3 to 9 yds/sec (8), and application of theory to various types of runs (6, 7). More extensive reviews of the literature will be found in these articles. The study to be described is concerned with cer-

[1]From the Department of Physical Education. The writer is indebted to his colleague Charles Keeney and to Coach Brutus Hamilton for aid in enlisting the subjects and for other aspects of co-operation that were important in securing the data. He is also deeply indebted to the men who did the running—all-out effort for 300 yards cannot be recommended as a recreational activity.

tain inter-relations of speed, velocity pattern and oxygen requirement, and extension of theory into regions that will permit practical application in sports activities.

Hypotheses

EQUATIONS OF MOTION

Acceleration Factor. It has been established (9) that the velocity of the runner at any time t, prior to the attainment of maximum velocity v_m, is given by the expression

$$dy/dt = v_m\ (1 - e^{-k_1 t}) \tag{1}$$

provided that he exerts himself maximally (or at least, steadily) up to the time of reaching top speed. It would be desirable to secure a more general mathematical expression of motion that would apply to the entire run, rather than just the initial phase. It must be recognized at the outset, however, that such an equation will only be applicable under specified conditions, as for example when constant effort is maintained throughout the run.

Fatigue Factor. It seems not unreasonable to start with the assumption that physiological fatigue begins at the very instant the runner starts down the track. For the present purpose this phenomenon can be considered in very broad terms—it might result from a decline in the biochemical release of energy, or from the accumulation of metabolites that indirectly decrease muscle energy. The net effect of this fatigue will be to reduce the speed potentiality of the runner with the passage of time. Tentatively, this situation can be described mathematically at any time t by some constant multiplied by the exponential term $e^{-k_2 t}$, where the exponent k_2 is a rate coefficient. If there is no fatigue, k_2 will be zero, and $e^{-k_2 t} = 1$. If k_2 is large, the magnitude of the term is unity at zero time, but declines rapidly (at a decelerating rate, however) with the passage of time.

Complete Equation. Using this exponential fatigue term to replace unity in the original Equation **1**, we may write

$$\frac{dy}{dt} = v_o\ (e^{-k_2 t} - e^{-k_1 t}) \tag{2}$$

to describe the velocity of the runner. The term v_o in the new equation is an "ideal" velocity that the runner would attain in the absence of cumulative fatigue; it may be defined as

$$v_o = v_m / (e^{-k_2 t_m} - e^{-k_1 t}\ ) \tag{3}$$

where v_m is the maximum velocity which he reaches at time t_m. The time t_m occurs when the *acceleration* or rate of change in velocity

$$\frac{d^2y}{dt^2} = v_o\ (k_1 e^{-k_1 t} - k_2 e^{-k_2 t}) \tag{4}$$

is equal to zero. Under the conditions of the problem, this time (approximately 6 seconds) is not particularly critical.

Equation 2 can be integrated by standard methods, giving

$$y = \frac{v_0}{k_2}(1 - e^{-k_2 t}) - \frac{v_0}{k_1}(1 - e^{-k_1 t}) \qquad (5)$$

where y is the distance attained by the runner at time t.

The above equations will not describe the acceleration, velocity, and position of a runner under any and all circumstances; it is indeed quite unreasonable to expect them to do so. There are however two special cases of particular interest where it can be hoped that they will prove valid—"all out" running, where *maximal effort* is maintained throughout the distance, and "even-pace" running, where submaximal *constant force* is maintained.

Oxygen Requirement. A theoretical equation for the oxygen cost of running has been presented in an earlier paper (8). Estimation of oxygen requirement by other methods and for an extended range affords an opportunity to test the equation. It will be convenient to develop this aspect of the problem in later sections of the report; the equation and its curves are shown in Figure I.

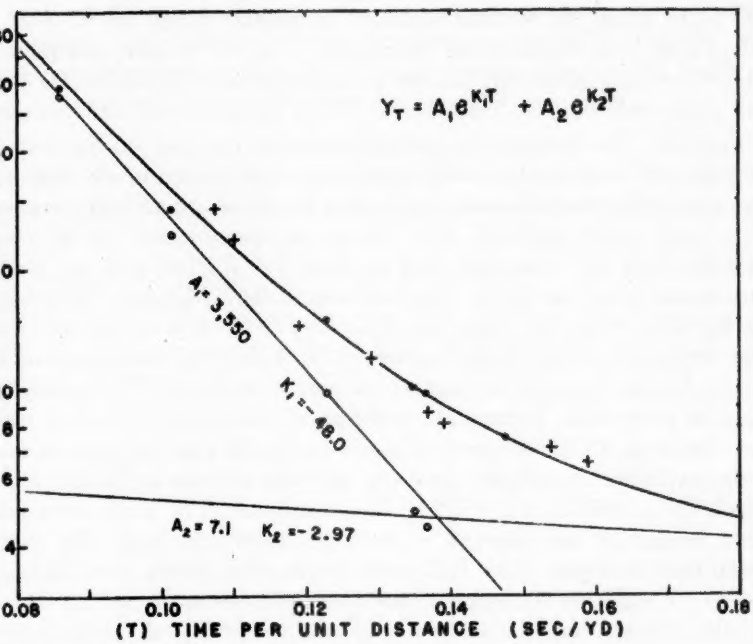
Methodology

Subjects. Data were secured on two groups of university men. In the first, there were 30 students from a professional major course in the theory and coaching of track and field sports. While some of these men had run competitively in high school, none had college experience. The second group consisted of 24 varsity dash men or middle distance runners. They were tested in the fall, so their times were not impressive. Even so, none of the first group ran as fast as the average of the second and none of the second was as slow as the average of the first.

Apparatus. A clap-board equipped with electrical contacts was used to start the run. It served to actuate one of the marking pens on a chronograph that could be read with an accuracy of approximately 0.01 seconds. Another pen was in series with contacts on the starting blocks, and also in series with contacts on timing gates placed at 10, 20, 40, 70 and 100 yards and thereafter at intervals of 40 yards up to a distance of 300 yards. The construction of these "gates" or timing stations has been described in other articles (7, 9).

Procedure. The men were tested individually, on a standard dirt track, with the first 140 yards on the straightaway. After a suitable warm-up, there was a 70-yard dash for the students and a 100-yard dash for the track squad. Following a rest, each man ran 300 yards with instructions to go as fast as possible throughout the run. A week or two later, the two runs were repeated. These tests were considered as a regular laboratory exercise for the students. The track men were volunteers who had evidenced interest in the tests, which were conducted in the afternoon during work-out hours.

Curve-Fitting. The k_2 coefficients have been obtained by plotting the log of mean velocity at each timing station against mean elapsed time, using semi-log graph paper. The intercept of the "decay" component with the zero time axis gives a value for v_0, while the slope of the line serves to estab-



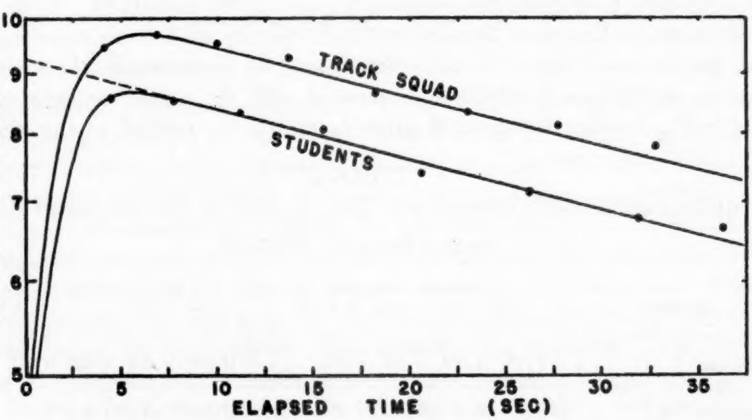


FIGURE. 1 (upper). Theoretical Curve of Oxygen Requirement as a Function of Speed. The crosses represent direct measurement, the solid dots were computed from champion times using the method of Table 3, and the open circles represent the oxygen debt component of the requirement estimated by that method. The straight-line function A_2—K_2 is the aerobic oxygen intake; it is summated vertically with the straight-line debt function A_1—K_1 to yield the curved line representing the theoretical oxygen requirement.

FIGURE II (lower). Theoretical and Observed Velocities during All-Out Running. The smooth curves represent the theoretical velocities; the plotted points are the velocities actually observed. Note that the ordinate of this figure is drawn to a logarithmic scale.

lish k_2 by using the method explained in another paper (8, *p. 171*). It is found that $k_2 = 0.00963$ for both groups, but v_o is quite different—10.42 yd/sec for the experienced runners compared with 9.22 yds/sec for the track and field students. The value $k_1 = 0.76$ is applicable to both groups.

Example. To illustrate the method, consider the data on *Students* shown in Figure II and Table I. The maximum velocity v_m in the 300 yd. run was apparently reached somewhat beyond the 40-yd. mark; interpolation on a large scale graph indicates 8.67 yds/sec as the amount. At 40 yards the time was 5.82 sec. Assuming that the peak was reached at 6 sec, it is necessary to add 1 yd. to the 40 (*i.e.* 0.18 sec. at 8.67 yds/sec). Substituting in the formula (9, *p. 413*, equation **5**) $t - y'/v = 6.00 - 41/8.67 = 1.27$. The reciprocal of this figure, namely 0.79, is the trial magnitude of k_1. To obtain k_2, the straight-line part of the curve in Figure II is extended leftward to zero time. Because the ordinate of this graph is in log units, the zero intercept (9.22 yds/sec) is equal to v_o. It may be seen on a larger scale graph that the velocity drops to half this amount in 72 sec., and $k_2 = 0.693/72 = 0.00963$ (8, *p.* 171). The magnitude of k_1 needs some readjustment because it was derived on a single-component basis. By trial, it is found that the figure $k_1 = 0.77$ gives the smallest errors (see Table 2). At the 70 yd. mark, it is observed that $t = 9.35$, hence $k_1 t = 7.2$ and $k_2 t = 0.0901$. Using a log table, $e^{k_1 t} = 1340$, and $e^{k_2 t} = 1.0945$, so the reciprocals (as demanded by the negative exponents) are 0.0007 and 0.914. Substituting these values in Equation **2** yields a 70-yd. velocity of 8.4 yds/sec compared with the observed figure of 8.5. Substituting in Equation **5**, the computed distance at 9.35 sec is 70.2 yds. compared with the actual distance of 70.0 (Table 2). Solving Equation **4** gives t_m as 5.8 sec instead of the assumed

TABLE 1

Observed and Predicted Distances and Curve-Constants in Dash and Longer Run

Physical Education Students

Measure	Separate 70 yd. run ($k = 0.80$ $v_m = 8.99$)					First 70 of 300 yd. run ($k = 0.80$ $v_m = 8.67$)				
Time (sec.)	2.15	3.40	5.65	9.04	....	2.19	3.48	5.82	9.35	....
Distance (yd.)	10.0	20.0	40.0	70.0	....	10.0	20.0	40.0	70.0	....
Predicted (yd.)	10.0	20.1	39.7	70.0	....	10.0	20.0	39.8	70.2	....

Track Squad

Measure	Separate 100 yd. run ($k = 0.79$ $v_m = 10.13$)					First 100 of 300 yd. run ($k = 0.79$ $v_m = 9.72$)				
Time (sec.)	1.98	3.14	5.20	8.17	11.22	2.04	3.25	5.38	8.44	11.55
Distance (yd.)	10.0	20.0	40.0	70.0	100.0	10.0	20.0	40.0	70.0	100.0
Predicted (yd.)	9.9	20.0	40.1	70.0	101.0	10.0	20.2	40.1	69.7	99.9

6 sec, and substituting in Equation **3** also checks out, since the difference between the two exponentials is 0.934 for either $t_m = 5.8$ or 6.0 min., and 9.22 yds/sec multiplied by this factor is 8.61—very close to the figure of 8.67 for v_m originally estimated from the graph.

TABLE 2

Observed and Predicted Distances and Velocities calculated from Observed Time Scores

Physical Education Students $v_o = 9.22$ yd/sec; $k^1 = 0.77$; $k^2 = 0.00963$)

Time	2.19	3.48	5.82	9.35	13.0	17.9	23.3	28.9	34.8	40.8
Distance	10	20	40	70	100	140	180	220	260	300
Predicted	10.0	20.1	40.2	70.2	100.1	139.1	180.1	221.5	260.5	299.0
Std. Error	0.15	0.27	0.46	0.8	1.2	2.2	2.0	2.5	2.8	3.0
Velocity	......	......	8.5	8.5	8.4	8.0	7.4	7.2	6.9	6.7
Predicted	......	......	8.6	8.4	8.3	7.9	7.5	7.2	6.8	6.4

Track Squad ($v_o = 10.42$ yd/ sec; $k^1 = 0.75$; $k^2 = 0.00963$)

Time	2.04	3.25	5.38	8.44	11.6	15.9	20.5	25.3	30.2	35.3
Distance	10	20	40	70	100	140	180	220	260	300
Predicted	9.9	20.3	40.4	70.7	100.5	139.3	180.0	220.1	259.1	298.0
Std. Error	0.13	0.23	0.46	0.8	1.1	1.6	1.9	2.3	2.6	2.8
Velocity	......	......	9.6	9.8	9.6	9.3	8.7	8.3	8.1	7.9
Predicted	......	......	9.6	9.7	9.5	9.2	8.8	8.4	8.0	7.6

Estimation of Oxygen Requirement. It was pointed out in the other reports that direct information on the oxygen requirement of running at speeds of competitive interest is not presently available (8) and that data for velocities above 6.5 yds/sec consist of observations on a single individual (11). Until direct information is available, the best that can be done is to estimate the requirement from world record times, assuming physiologically maximum magnitudes of 5.5 L/min for oxygen income and 18 liters debt. Table 3 has been prepared for this purpose. The figures are net, in that no correction has been made for the "pull-up" after crossing the finish line. The income has been computed on the basis of the formula presented in another article (5), thus correcting approximately for the time required for the rate of intake to build up to the steady-state value. A theoretical correction has also been applied to the velocity, in order to account for the time lost during the acceleration phase of the run (9).

The 100-yard figures are shown in the table as a matter of interest, but are not usable because the amount of debt possible in such a short period (less than one circulation time) is quite uncertain, although surely less than the maximal value. There is also considerable uncertainty about the true amount of oxygen income for the 220 and 440, but the total error from this cause is small because most of requirement is handled by debt. It may be

mentioned that recent changes in record times, as for example the new figure of 1:48.6 for the 880, do not alter the estimations to any important degree.

The oxygen requirements from Table 3 have been plotted as solid dots in Figure I. The points in this figure indicated by crosses are the oxygen requirements measured on a single subject by Sargent (11) in 1926. The smooth curve is a plot of the equation for oxygen requirement

$$Y = A_1e^{K_1T} + A_2e^{K_2T} \tag{6}$$

using the curve constants derived in the previous study, with the exception of a slight decrease in A_1 that improves the predictive accuracy of the formula at high speeds. This change is not sufficient to appreciably alter the curve published as Figure II of the other article (except at the very highest points where data are sparse) so that the present computations tie in rather precisely with the known oxygen requirement at lower velocities (8).

TABLE 3

Oxygen Requirement of Running computed from 1952 World Records

Assuming 18 liters maximum debt and 5.5 liters/min. maximum intake

Distance (yd.)	Time (min:sec)	Av. v_m (yd/sec)	v_m (yd/sec)	O_2 Debt (liters)	O_2 Income (liters)	Total O_2 (liters)	Debt (%)	O_2 Requirement (liters/min)
100	0:09.2	10.87	12.50	12.6?	....	12.6?	....	83.0?
220	0:20.2	10.89	11.57	18.0	1.7	19.7	91.5	58.5
440	0:46.0	9.57	9.82	18.0	3.0	21.0	85.7	28.2
880	1:49.2	8.06	8.14	18.0	9.3	27.3	66.1	15.0
1640	3:43.0	7.36	7.40	18.0	20.5	38.5	46.8	10.35
1760	4:01.4	7.29	7.33	18.0	22.1	40.1	44.9	9.95
3520	8:40.4	6.76	6.78	18.0	47.7	65.7	27.4	7.58

Results and Discussion

VELOCITY AND DISTANCE PREDICTIONS

Validation of Formulae. The timing stations were not close enough together in the first 50 yards to permit measuring the velocities in this part of the run. Using the mid-points of the other stations as reference points, the velocities have been measured and compared with the values computed from the formula. It may be seen in Figure II that the speed pattern follows the trend of the theoretical curve, the low points at 18 and 21 seconds presumably representing the consequence of rounding the track curve which was entered at the distance 140 yards. There is a definite tendency for the velocity near the end of the run to be somewhat higher than predicted. This may be a typical "end-spurt" of psychogenic origin or it may mean that the runners are approaching a steady-state asymptotic velocity that could be maintained indefinitely. The experiment would have to be repeated using a longer distance in order to clear up this point. A more exacting check is obtained by substituting the curve constants in Formula 4, thus comparing the theoretical and observed distances at the observed times at every station (Table 2). The standard errors have been computed by converting the time

deviations into equivalent distances on the basis of the velocity at each station. It can be seen by inspection that the sampling errors are larger than the errors of prediction.

Variation of Velocity as a Function of Distance. Since Equation **2** (velocity) and Equation **5** (distance) are comparable exponential functions involving the variable time t, it follows that velocity is linearly related to distance at any time t after peak velocity is attained (see Equation **4**). In other words, the drop-off in velocity is constant for each unit of distance. Using the experimentally determined k's and substituting appropriate values of v_o to secure the desired race times, it can be shown that the "all-out" run has the same velocity profile as the plan recently recommended by Doherty for the 440 yard event (4). His 880 yard plan has a similar pattern, but with somewhat less drop-off, corresponding to $k_2 = 0.0013$. For purposes of computation, it is mathematically more convenient to handle the velocity profile as a function of elapsed time rather than distance.

Maximal vs. *Sub-Maximal Effort.* Even though the men were instructed to run with maximal speed throughout the run, it cannot be assumed that they did so. For this reason, it is of interest to examine the separate 70 and 100 yard dashes in comparison with the corresponding parts of the 300 yard runs. It is necessary to use the simple one-component method (Formula **2** of reference 9), namely

$$y = v_m \left(t - 1/k\, e^{-kt} - 1/k\right) \tag{7}$$

because the k_2 component cannot be determined in the separate short sprints. The curve coefficients have been computed and are given in Table 1. The time scores under the two conditions show that the longer run was done at somewhat less than maximal effort. Nevertheless, the formula is applicable for both conditions. The maximum velocity v_m varies in proportion to the speed of the run, whereas the k's are unaltered, and moreover are almost identical in the two different groups covering a velocity range of 16.5 per cent. This confirms the previous report that fast and slow runners do not differ with respect to their k_1 coefficients (9).

Even though unrelated to speed at distances of 70 or 100 yards, the k's exhibit individual differences. The test-retest correlation is $r = 0.66$ for the sprint, and 0.42 for the longer run—somewhat less than the figure of $r = 0.75$ found in the previous study (using a more adequate method with timing stations at 5-yard intervals) but statistically significant without question. The v_m data are confirmed almost exactly—8.99 yds/sec for the students (in the separate sprint) compared with 8.9 in the other report, and $r = 0.93$ (students) or 0.87 (track-squad) compared with the reported v_m reliability of $r = 0.91$ (9).

Steady Pace Running. It should be noted that the k_1 coefficients obtained by the simple formula are 4.6 per cent larger than when the same data are fitted with the two component system. This is an inevitable mathematical result of the slight increase in curvature resulting from the influence of the second component, and is of no theoretical consequence. The rather re-

markable stability of k_1 over a considerable range of v_m values justifies the use of Formula 7 and derived expressions such as Formula 1 as equations of motion for steady-paced running. (It will be recalled that the relations of these simpler formulae to the mathematical description of "all-out" running were developed in the introduction.) The preceding section has established the validity of the equations for the acceleration phase of the run, and they are necessarily accurate thereafter if a constant velocity is maintained.

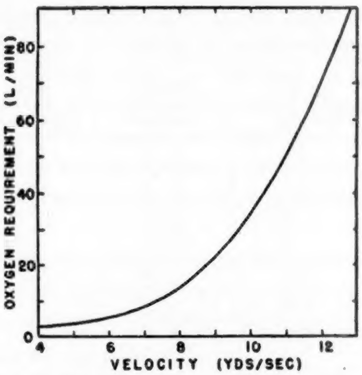
Comparative Efficiency of Running "All-Out" or "Steady Pace." Using the oxygen requirement data of Figure I and the appropriate equations of motion, it is a straightforward (although tedious) process to compute the requirement for successive short periods of elapsed time and summate them to estimate the amount of oxygen required for a run of specified speed and velocity profile. In the comparisons that will be made, the steady-pace profile has been modified to provide that the runner will, during the last 20 yards of the race, decelerate from the standard velocity to the same terminal velocity at the finish line that characterizes "all-out" running and the Doherty-type plan. This provision equates the runs with respect to total elapsed time, and also with respect to residual energy as estimated by final velocity at the tape. The first 2½ seconds of the runs have also been equated to avoid any controversy as to uncertainties concerning the oxygen requirement of this part of the race.

Two examples have been worked out. In the first, the Doherty-plan 440, run in 46 seconds (4), is found to involve a peak velocity v_m of 11.40 yds/sec and a terminal velocity of 7.91 yds/sec, with an oxygen cost of 25.30 liters. For the same race time, the steady-pace profile reaches a v_m of 9.90 yds/sec and requires 24.87 liters of oxygen. Using a method that will be illustrated in the next section, it can be shown that with the oxygen cost of both plans equated at 25.30 liters, the steady-pace method would result in 0.2 seconds faster time. Thus there is no serious physiological objection to the Doherty plan for the 440 *when run this fast,* and no reason to deny the other advantages that have led him to advocate it.

In the second example, a Doherty-plan 880 would be run in 1:50 with v_m at 8.60 yds/sec, a terminal velocity of 7.51 yds/sec, and 26.275 liters of oxygen used. The steady-pace plan would have $v_m = 8.01$ and require only 25.00 liters. With equated oxygen cost, the steady-pace profile would save 1.65 seconds. It is surprising that the 440, run at a considerably higher velocity than the 880, has a *smaller* per cent advantage in time for the steady-pace profile than the half mile. The reason for this result will be explained below. It would seem, in the light of this example, that there is a strong physiological reason for objecting to variable pace in the 880.

Influence of Velocity on Advantage of Steady-Pace. Figure III, relating oxygen requirement to velocity using linear co-ordinates, shows a decided "knee" or region of maximum curvature. (This knee is even more striking when requirement is plotted against time per unit distance, i.e. the reciprocal of velocity.) At velocities well below or well above the knee, a specified

change in velocity does not greatly influence the efficiency of running. In contrast (for velocities *within* the knee and more particularly just at its greatest curvature) the oxygen cost of running a specified amount slower than some reference speed does not greatly decrease, whereas the same amount of velocity change *in the direction of increased speed* results in a disproportionately large increase in cost. A variable pace is therefore much less efficient than a steady pace.



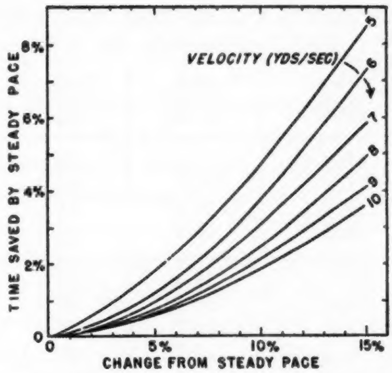


FIGURE III (left). Oxygen Requirement *vs.* Velocity. This is an arithmetic plot of the data of Figure I.

FIGURE IV (right). Efficiency *vs.* Speed Variation. The curves show the gain in efficiency as a function of the amount of variation in pace, for several speeds.

The consequence of variable-pace running can be visualized by computing the influence of a standard *per cent* change in speed for a series of velocities, stating the results as *per cent* time saved by the steady-pace profile. The method of doing this can best be explained by an example. In one minute, a distance of exactly 360 yards would be covered by running steady-pace at 6.00 yds/sec. The same distance would be covered by running a half-minute at 6.6 yds/sec at a cost of 3.500 liters of oxygen and a half-minute at 5.4 yds/sec at a cost of 2.315 liters. With a $\pm$ 10-per-cent variation in pace, the cost per minute is therefore 5.815 liters. It would be possible to run the 360 yards at 6.26 yds/sec in 57.5 seconds at a cost of 6.085 L/min or 5.83 liters for the elapsed time—but this turns out to be somewhat too fast and too costly. For a second try, the 360 yards could be run in 57.8 seconds at a speed of 6.23 yds/sec and a cost of 6.00 L/min or 5.78 liters for the exact distance—slightly too slow. By interpolation, it is seen that a speed of 6.25 yds/sec should be correct. At this speed the distance can be run in 57.6 seconds at a cost of 6:055 L/min or 5.815 liters for the 360 yards. Thus the oxygen cost is exactly the same to run this distance in 60.0 seconds at a speed of 6.0 yds/sec $\pm$ 10 per cent or in 57.6 seconds at a steady pace of 6.25 yds/sec, and the time saved by the latter profile is 2.4/60 or 4.0 per cent.

Table 4, computed by the above method, shows that the greatest increase in efficiency due to steady-pace running is in the anticipated region of speed,

namely 7 to 8 yds/sec. The region of greatest time saved is considerably lower, namely 5 to 6 yds/sec. No doubt this depression of the optimium region is caused by the successively lower efficiency of running that becomes increasingly and rapidly evident as velocity is increased—the oxygen saved by steady-paced running is of progressively less value because any gain in speed is increasingly costly.

TABLE 4

Time Saved by Running Steady-Pace, using the Same Oxygen Requirement as Running Irregularly with ± 10-per-cent Variation in Speed

Average Velocity (yd/sec)	Steady O_2 Cost (L/min)	Irregular O_2 Cost (L/min)	Increase in O_2 Cost (per cent)	Possible Velocity (yd/sec)	Time Saved (per cent)
12.0	70.6	73.6	4.2	12.2	1.5
11.5	59.7	62.6	4.6	11.7	1.6
11.0	50.3	53.0	5.0	11.2	1.7
10.5	41.9	44.1	5.2	10.7	1.8
10.0	34.4	36.3	5.5	10.2	1.9
9.5	28.0	29.9	6.2	9.7	2.1
9.0	22.2	23.6	6.7	9.2	2.2
8.5	17.5	18.9	7.5	8.7	2.3
8.0	13.6	14.7	7.9	8.2	2.5
7.5	10.7	11.4	7.7	7.7	2.7
7.0	8.3	9.0	8.0	7.2	3.1
6.5	6.7	7.1	6.5	6.7	3.4
6.0	5.5	5.8	5.2	6.3	3.8
5.5	4.7	4.9	2.7	5.7	4.1
5.0	4.2	4.3	1.8	5.2	4.0
4.5	3.7	3.8	0.9	4.6	1.4
4.0	3.4	3.4	0.3	4.1	0.2

Using the same method, the family of curves in Figure IV has been computed to show the amount of gain in efficiency by running steady pace, as a function of the amount of irregularity in velocity pattern. Since the curves are concave upward, the larger degrees of irregularity cause a relatively greater amount of gain. A corresponding family of curves (not shown) giving amount of time saved as a function of linear (rather than per cent) changes in velocity, is found to be quite similar to the curve-family illustrated. It should be a simple matter, with the aid of Figure IV and Table 3, to estimate the quantitative influence of variable pace in practical problems. The loss of efficiency is greatest at high speeds, but the loss in efficiency *due to irregular pace* is greatest at low speeds (except that the trend reverses at extremely low speeds). The distance runs are therefore more sensitive to pacing for two reasons: The effect is greater for runs of this type, and it cumulates for a longer period of time. Nevertheless, the effect is large enough to justify consideration in half-mile runs, and even for the quarter when run

at typical high school (as contrasted with national record) velocities. It must be recognized, of course, that other factors in addition to physiological efficiency are involved in competitive running(4).

Previous analyses of the greater efficiency of steady-pace running made by the writer and by others (6) have assumed the validity of the empirical "power law"

$$O_2\ Requirement = Cv^{3.8} \tag{8}$$

widely printed in text books. This mathematical expression is characterized by greater curvature for low velocities and a less well defined "knee" than the function employed for the present calculations. The formula used for Figure I was developed from a theoretical analysis of the directly measured oxygen requirement of running at various speeds (8), and the knee of the curve as well as its lower limb were established by those data without reference to estimations based on world record performances as used in the present study.

Confirmation of Oxygen Requirement Theory. Figure I is a magnified and left-ward extended version of Figure II of the earlier study on oxygen requirement (8), using the same mathematical function with the relatively minor readjustment of the A_1K_1 intercept as mentioned in a previous section. The fact that the curve fits the points computed from record times justifies considerable respect for the formula. A more valid confirmation may be secured by computing (from Table 3) the amount of the oxygen requirement that is composed of oxygen debt. The A_1K_1 component was derived hypothetically as representing this factor whereas the A_2K_2 component represented the aerobic factor (8, *p. 170*). It may be seen in Figure I that the debt factor (plotted by open circles) does indeed tend to fall on the A_1K_1 component. This is a finding of considerable interest with respect to explaining the inefficiency of fast running (see for example references 1, 2, and 3).[2]

Some uncertainty is of course still present concerning the exact magnitude of the curve coefficients, and will remain until adequate direct information on oxygen requirement is available at the higher velocities. It should be mentioned that a direct resolution of the A_2K_2 component will probably never be experimentally possible at the higher velocities because of technical difficulties arising from the necessarily short period of observation. As mentioned earlier, the aerobic intake rate must decline if there is insufficient time for the oxidative process to build up (5).

---

[2]It must be realized that the inefficiency probably has several causes. For example, when muscular work was done at two speeds of movement on the bicycle ergometer, with the load adjusted so that the metabolic cost was the same in both cases, the mechanical efficiency of the faster movement was considerably less (Henry, F. M., Janice De Moor and I. R. Trafton, *Research Quarterly* **22**: 324-333, 1951). Computations of gross efficiency based on the data of E. Asmussen (*Acta Physiol. Skand.* **28**: 364-382, 1953), indicate however that this effect is largest for light work loads and probably less important for the current problem.

Empirical Formulae for Oxygen Requirement. In the previous study (8), two empirical formulae were presented for the purpose of computing oxygen requirement from speed, and speed from oxygen available. While valid for the purpose when used within the original velocity range, neither of these formulae can be successfully extended to the higher velocities considered in the present paper. This is typical of the limitations of empirical methods.

Reaction Time

Type of Data. The 24 varsity runners were given 50 trials with each hand, and 50 with each foot, reacting to a sound stimulus. These scores were correlated with the average of the two 100-yd.-dash runs. It should be noted that there were reliable individual differences in sprint time ($r = 0.91$, $S.D. = 0.56$ sec.)

Absence of Correlation. Using the pooled hand reactions, the correlation with sprint time is $r = -0.094$. Using the foot reactions, it is $r = 0.038$. This finding agrees with a number of studies (6, and 9, *p. 419*, as well as others from this laboratory). It can be concluded unequivocally that, at this university, there is no relationship between speed of motion and speed of reaction in sprinters or in nominally "unselected" groups. There remains the possibility that sprinters, as a group, may have somewhat faster average reaction times than certain other selected groups of athletes.

Summary and Conclusions

An equation of motion for all-out running was derived by postulating an exponential fatigue factor that subtractively reduces the steady-state asymptote of a previously validated equation describing velocity and distance during the acceleration phase of a run. It was found that the distance reached at any particular time t after the start of a 300-yard all-out run could be predicted within 1-per-cent error by this equation of motion.

Data from a group of 24 varsity runners yielded k_1 and k_2 coefficients numerically the same as from data on 30 physical education major students inexperienced in running. The 16-per-cent differential in running time between the two groups must therefore be attributed to differences in the velocity factor v_o. Individuals of both groups were also tested in a short dash (70 or 100 yards), which was run somewhat faster than the corresponding part of the longer distance. Since the k_1 coefficients were unaltered by speed in this situation also, it was concluded that they could be used in the equation of motion for steady-pace running. For this type of velocity pattern, the coefficient k_2 becomes zero.

In order to examine the practical consequences of running steady-pace compared with the declining velocity pattern currently advocated by some experts, or with other variable-pace race patterns, it was necessary to extend oxygen requirements theory to velocities for which no direct data were available. Computed requirements based on world record race times and physiologically reasonable magnitudes of maximal oxygen debt and intake were found to be consistent with a previously derived mathematical statement of

the theory. Particular importance was attached to the close agreement between the magnitude of the oxygen debt component as a function of speed as computed by this method and as predicted by theory. This result suggested that the inefficiency of running at fast speeds is to a considerable extent a result of the increasingly large component of anaerobic metabolism.

The relatively greater physiological economy of steady-pace running, as estimated by this theory, is highest for velocities within the range of 7 to 8 yds/sec but still quite evident at 9 or 9.5 yds/sec. The maximum amount of time that can be saved, however, is in the velocity region 5 to 6 yds/sec and declines progressively at higher velocities. This difference is a result of the rapidly increasing oxygen cost of increasing velocity, causing the oxygen dividend of steady-pace running to be of progressively less value in permitting greater speed.

Nevertheless, the advantage of the steady pace is theoretically large enough to be considered of practical importance for distances as short as the half mile. For example, a 1:50 half, if run according to a recommended variable-pace plan, would cost 26 liters of oxygen—sufficient to run the distance in 1:48.3 by the steady-pace method. A roughly similar amount of time would be saved in a typical high school 440, but an Olympic record time for the latter distance would be improved only about 0.2 seconds. Tables and curves are given to simplify application of the theory to other specific practical situations.

The varsity runners were tested as to speed of hand reaction and foot reaction to a sound stimulus. In agreement with several other recent studies, the correlation between reaction time and sprint speed was found to approximate zero.

REFERENCES

1. Christensen, E. H., and P. Högberg. The efficiency of anaerobical work. *Arbeitsphysiol.*, **14**: 249-250 (1950).
2. ———, Steady-state, O_2-deficit and O_2-debt at severe work. *Arbeitsphysiol.*, **14**: 251-254 (1950).
3. De Moor, Janice, Individual differences in oxygen debt curves related to mechanical efficiency and sex. *J. Applied Physiol.*, **6**: 460-66 (1954).
4. Doherty, J. K., *Modern Track and Field.* New York: Prentice-Hall, (1953).
5. Henry, F. M., Aerobic oxygen consumption and alactic debt in muscular work. *J. Applied Physiol.*, **3**: 427-438 (1951).
6. ———, Research on sprint running, *Athletic J.*, **32**: 30-34, (Feb. 1952).
7. ———, Force-time characteristics of the sprint start. *Research Quarterly*, **23**: 301-318 (1952).
8. ———, The oxygen requirement of walking and running, *Research Quarterly*, **24**: 169-175 (1953).
9. ———, and I. R. Trafton, The velocity of sprint running, with some observations on the muscle viscosity factor. *Research Quarterly*, **22**: 409-422 (1951).
10. Hill, A. V., *Muscular Movement in Man*, New York: McGraw-Hill (1927).
11. Sargent, R. M., The relation between oxygen requirement and speed in running. *Proc. Roy. Soc.*, (B) **100**: 10-22 (1926).

(Submitted 8-17-53)

Minimum Muscular Fitness Tests in School Children

HANS KRAUS, M.D.
RUTH P. HIRSCHLAND

Institute of Physical Medicine and Rehabilitation
New York University—Belevue Medical Center
New York, New York

Abstract

Six test movements appraising strength and flexibility of trunk and leg muscles were given to 4264 American and 2870 European children from comparable urban and suburban communities. 57.9% of the Americans failed and 8.7% of the Europeans. The poor American showing can be explained by our high degree of mechanization obviating much physical activity. Since previous studies have shown that these tests represent minimum muscular fitness, and that falling below these levels predisposes to orthopedic and emotional difficulties, it is urged that the physical activities of our children be increased and that muscle tests be given at regular intervals, and made a part of the child's complete school record, to assure at least these minimum standards for our children.

THE CONNECTION between good posture and health has for many years been an established part of physical education and health education. In the last decades, the concept of posture has shifted more and more from merely structural to a functional one as well. The interest in posture has gained added impetus because of the attention paid to flexibility in the present day treatment of poliomyelitis.

Studies on the subject were published by A. D. Gurewitsch and Margaret A. O'Neill in 1944 (4) and by H. O. Kendall and Florence P. Kendall in 1948 (5).

In 1945 Hans Kraus and S. Eisenmenger-Weber reported a follow-up study on 200 posture cases (school children) who had been patients at the posture clinic of Columbia Presbyterian Medical Center (12, 13, 14).

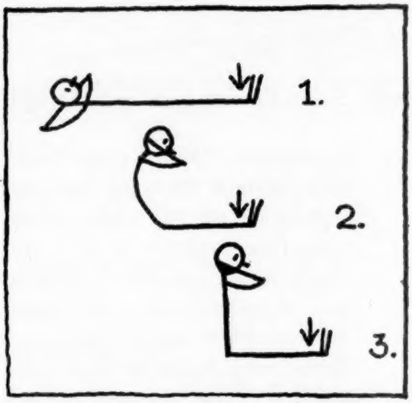
The results reported in the above three papers were evaluated, not only by structural measurements, but by functional measurements as well. These functional measurements included strength tests for trunk muscles and flexibility tests for trunk and hamstring muscles.

The key tests were adopted as basic tests in a Low Back Clinic which was organized by Dr. Barbara Stimson at Columbia Presbyterian Hospital under the auspices of Dr. William Darrach in 1946.

The Kraus-Weber Tests for Muscular Fitness are not designed to determine optimum levels of muscular fitness, but rather to determine whether or not the individual has sufficient strength and flexibility in the parts of the body upon which demands are made in normal daily living.

Kraus-Weber Tests for Muscular Fitness

(*There should not be any warm-up prior to taking the tests.*)



TEST 1.

Purpose: Tests the strength of the abdominals and psoas.

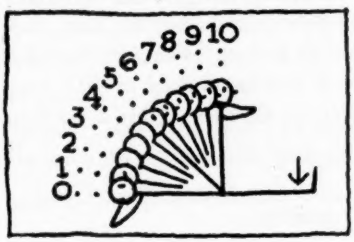
Designation: "Abdominals plus psoas" or *A+*.

Position of Person Being Tested: Lying supine, hands behind neck. The examiner holds his feet down on the table.

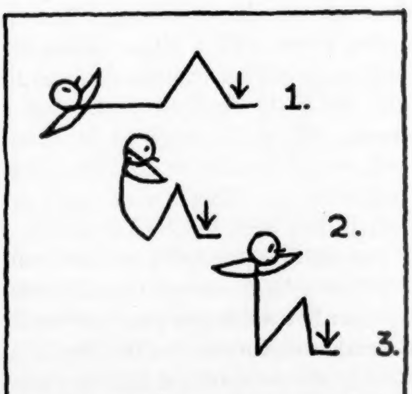
Command: "Keep your hands behind your neck and *try to roll up* into a sitting position."

Precaution: If the person being tested is unable to perform this movement at first try, it may be because he has not understood the directions. Help him a little and then let him try again. Watch for a "stiff back sit-up." This may indicate that either he has not understood you and needs a further explanation with emphasis on "rolling up," or that he has *very* poor abdominals and is doing most of the work with his psoas.

Watch also for a twist of the upper body as he sits up. This may be due to unequal development of the back muscles.



Marking: If the person being tested cannot raise his shoulders from the table, the mark is *0*. If unaided, he is able to reach a sitting position, the mark is *10*. If the examiner must help him halfway to the sitting position, the mark would be *5*. The distance from supine to sitting is marked from *0* to *10*.



TEST 2.

Purpose: Further test for abdominals.

Designation: "Abdominals minus psoas" or *A—*.

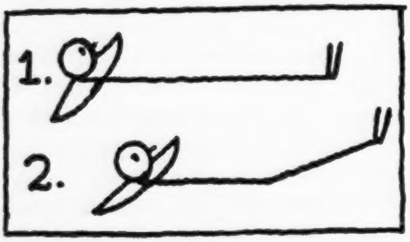
Position of Person Being Tested: Lying supine, hands behind neck and knees bent. Examiner holds his feet down on the table.

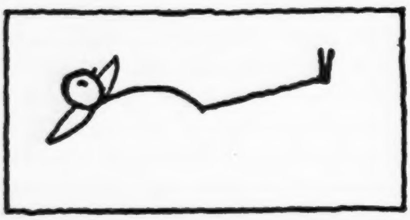
Command: "Keep your hands behind your neck and *try to roll up* into a sitting position."

Precaution: The precautions are the same as for Test 1, but as Test 2 is

usually more difficult the tendency toward "stiff back sit up" will be even more pronounced and to it is added the tendency to help with one or the other elbow.

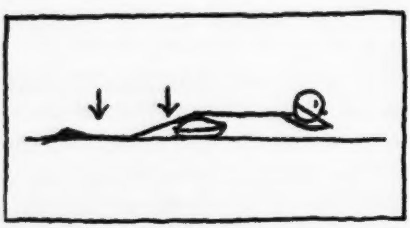
TEST 3.
Purpose: Tests the strength of the psoas, and lower abdominals.
Designation: "Psoas" or *P.*
Position of Person Being Tested: Supine with hands behind neck and the legs extended.





Command: "Keep your knees straight and lift your feet ten inches off of the table. Keep them there while I count." The count is ten seconds. (Adding any three syllable word after each number makes the count fairly reliable as to time. For example, "One chimpanzee, two chimpanzee, three chimpanzee, etc.")
Precaution: If the person tested has not understood your command, he may try to raise his chest when he raises his feet and will need further explanation. Watch for an extremely arched back which may indicate very weak abdominals or postural habits contributing to sway back or lordosis.
Marking: Holding for ten full seconds is passing and is marked as *10.* Anything less is recorded as that part of the ten seconds that was held: *4* for four seconds, or *7* for seven seconds, etc.

TEST 4.
Purpose: Tests the strength of the upper back muscles.
Designation: "Upper back" or"UB."



Position of Person Being Tested: Lying prone with a pillow under his abdomen, but far enough down as to give the body the feeling of being a seesaw which, if weighted at either end, would be able to hold the other end in the air. This is most easily accomplished with these commands:
Command: "Roll over onto your stomach and lift up the middle so that I can slide this pillow under you." (Be sure the pillow is large enough to really support him.) "Now, I am going to hold down your feet while you put your hands behind your neck and raise up your chest, head, and shoulders. Hold them up while I count." The count is ten seconds.

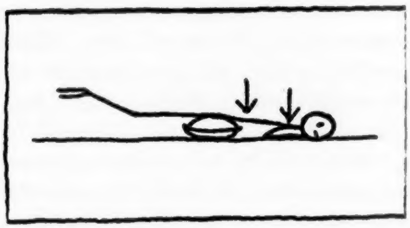
Precaution: Do not let the person being tested drop his chest onto the table or rest his elbows. Watch for pronounced muscular development on one side of the spine. If this condition is present, the back should be checked from time to time to guard against scoliosis (curvature ofthe spine).

Marking: Holding for ten full seconds is passing and is marked as *10.* Anything less than ten seconds is recorded as that part of ten seconds that was held. For example, a person staying up for four seconds would get the mark of *4.*

TEST 5.

Purpose: Tests the strength of the lower back.

Designation: "Lower back" or *LB.*



Position of Person Being Tested: He remains prone over the pillow, but removes his hands from behind his neck, places them down on the table and rests his head on them.

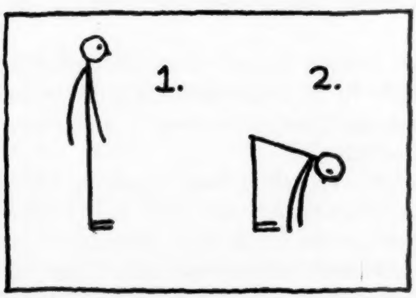
Command: "I am going to hold your chest down on the table; try to lift your legs up, but do not bend your knees." There may be a tendency to bend the knees or even to support the legs by keeping the toes on the table. It may be necessary to assist him to the required position. "Now, hold this position white I count." The count is ten.

Marking: Holding for ten full seconds is passing and is marked as *10.* Anything less is recorded as that part of the ten seconds that was held, for example, four seconds would be *4.*

TEST 6.

Purpose: Tests the length of back and hamstring muscles.

Designation: "Back and Hamstrings" or *BH.*



Position of Person Being Tested: Standing erect in stocking or bare feet, hands at his sides.

Command: "Put your feet together, keep your knees straight, now lean down slowly and see how close you can come to touching the floor with your finger tips. Stay down as far as you can for a count of three. DO NOT BOUNCE.

Precaution: Watch out for bouncing. The furthest point reached without bouncing and held for three counts is the marking point. The examiner should hold the knees of the person being tested in order to prevent any bend.

Marking: Touch is designated with *T. Touch* is only given when the floor-touch is held for three counts. Less than *Touch* is marked by the distance in

inches between the floor and the finer tips. For example, a person unable to touch the floor by two inches would be marked, *—2"*.

In the above tests, the words "upper" and "lower" are merely used to indicate test movements rather than any specific areas.

Discussion

Over 4,000 cases of patients with low back pain were thus evaluated. The people were drawn from the Low Back Clinic at Columbia Presbyterian Hospital and from the Low Back Clinic of New York University's Institute for Rehabilitation and Physical Medicine, as well as from private practice.

A high percentage of these cases (approximately 80%) (19), after thorough evaluation by a team of specialists, were found free from organic disease; however, this group failed to pass one or more of the above tests. When treated with therapeutic exercise, they improved as their test results improved. In an 8-year follow-up, it was found that as the patients stopped exercising, their tests failed and their complaints reappeared. This condition could be reversed by resuming either the therapeutic exercise or regular physical activities, but the final outcome—and the permanency of relief—ran parallel with the muscle status. (6, 7, 8, 9). These impressions are shared by other authors (17, 18, 19, 20, 21).

These tests represented *minimum-fitness tests;* that is, they were tests which indicated a level of strength and flexibility in certain key muscular groups below which functioning of the whole body as a healthy organism seemed to be endangered. Furthermore, patients whose physical fitness level fell below these *minimum requirements* appeared to be "sick people," individuals who bore all the earmarks of "constant strain" (Selye, 16), and who frequently manifested signs of emotional instability.

This fact seems quite understandable in the light of W. B. Cannon's study on bodily reactions to physical threat or fear, which emphasized the necessity to discharge those energies accumulated as "preparation for flight or for flight" (1).

When the individual is prevented from ridding himself of surplus energies, tension remains. This tension often results in physical discomfort. It is this observation which is of particular interest to those attempting to understand and alleviate muscular pain of obscure origin (15).

Since the percentage of backache sufferers—as well as that of patients suffering from other tensions syndromes—was undoubtedly very high, the question arose as to whether these syndromes were not the result of an imbalance: an imbalance caused by excessive intake of stimuli and by too little outlet, especially too little physical outlet. If this were the case, prevention might well be found in an increase of physical activity (7, 10, 11).

In a first step toward prevention of this disease, it seems logical to develop healthy movement habits in children, as well as attempting to develop their bodies above and beyond minimum standards of flexibility and strength. Children will not immediately feel the impact of relative muscular weakness and

stiffness; but as they proceed in life, and as they must draw on reserves and muscle potential, they shall have to fall back on what was formed during their early years. If high muscular standards are present in the first two decades of life, maintaining these standards, or regaining them later on, will be much easier.

When we started to determine the muscular fitness levels of American school children by submitting them to the above tests, we found such a high percentage of failure that we felt it necessary to make comparative studies abroad.

The question arose in our minds whether our minimum standards were not too high to be met by the average school child. Over 10,000 school children have been tested to date both here and abroad. A report on over 4,000 American children and 2,000 European children (Italian and Austrian) was published in the *Journal of the American Association for Health-Physical Education-Recreation* in December 1953 (11). We here submit a more complete study, augmented by information gained from testing 1,156 Swiss children.

The children reported in this study were between the ages of 6 and 16 years, and were from public school systems in suburban and small urban communities. The sizes of these cities, both European and American, were comparable. Every effort was made to keep test conditions identical and to make the tests uniform. All the tests were made by the authors themselves and were completely standardized. Statistics obtained from sick or disabled children were excluded from the results, although all the children on hand were tested in order to avoid trauma to disabled individuals.

The major difference between these two groups is the fact that European children do not have the "benefit" of a highly mechanized society; they do not use cars, school-buses, elevators, or other labor-saving devices. They must walk everywhere—even to school, frequently a long distance. Their recreation is largely based on the active use of their bodies. In this country the children are generally conveyed in private cars, or by bus, and they engage in recreation as spectators rather than as participants.

Test Results

The following table gives the results of the tests:

	Austrian	*Italian*	*Swiss*	*American*
Number Tested	678	1036	1156	4264
Failure	9.5%	8.0%	8.8%	57.9%
Incidence of Failure	9.7%	8.5%	8.9%	80.0%

The Kendalls reported a very definite variation of results with age (5) and we fully agree with the authors that the ages between 10 and 13 years seem to be "critical" ones. We found, however, that this "critical" period was reached one to two years earlier in Europe.

It is interesting to note that European figures, representing completely different countries, are highly similar (Figure I). As these countries are effectively separated by frontiers, we consider this of significance. The higher flexibility failure rate in Austrian and Swiss children, as compared with

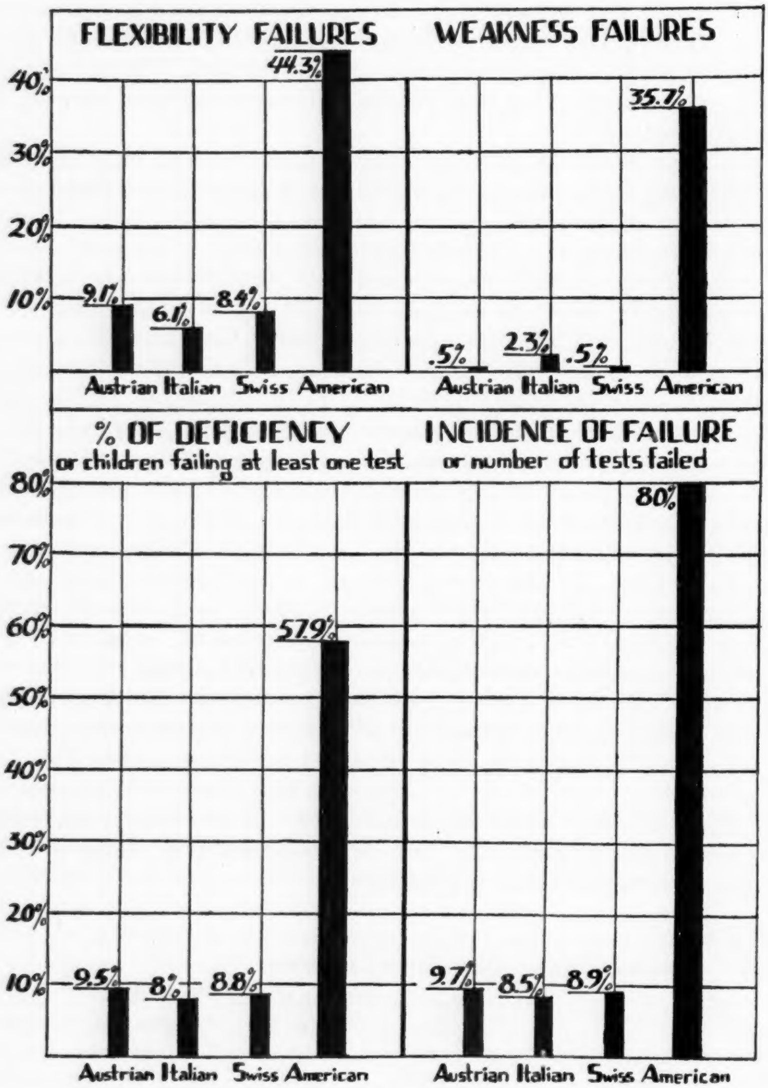


FIGURE I. Comparison of Test Failures in Different Countries.

that in Italian children, is due partly to the fact that not much attention is paid to flexibility in an otherwise strict and vigorous physical education program. A one-sided strengthening program, disregarding flexibility and relaxation, could well result in shifts of these figures.

Findings

Figures II and III concern the findings on the progress of muscle deficiency during the school years (6 to 16).

In our efforts to determine the age at which muscle deficiency first becomes apparent, we were met with a distressing fact: children coming into the first grades of the school system are already seriously deficient.

Furthermore, it appears that we are unable to alleviate this situation during the time the children are in elementary schools. They leave elementary school in very much the same condition as when they entered it—if anything, a little worse.

Weakness as well as flexibility failures show that at no time do American statistics approach the fitness levels of the European (Figure IV).

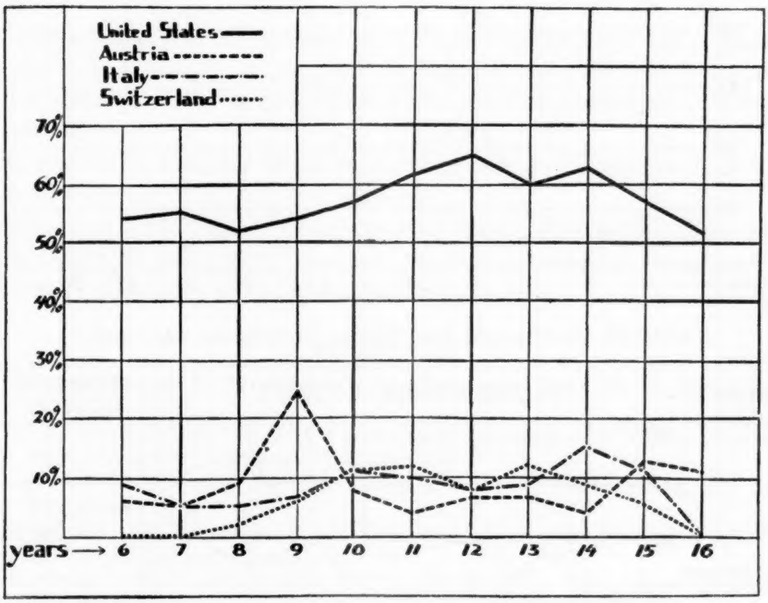


FIGURE II. Percentage of Test Failures at Different Age Levels.

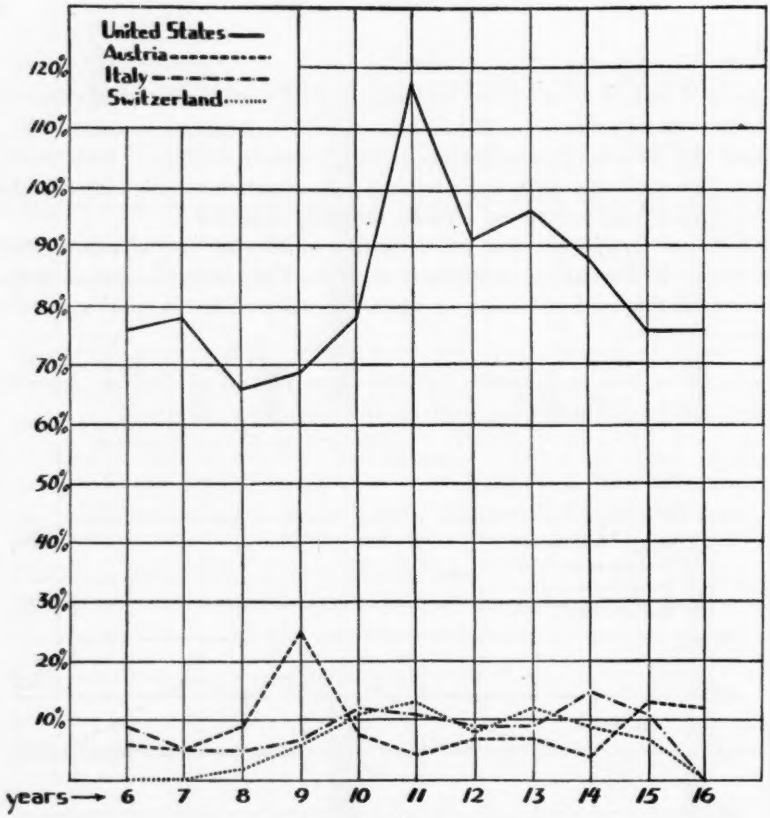


FIGURE III. Incidence of Test Failures at Different Age Levels.

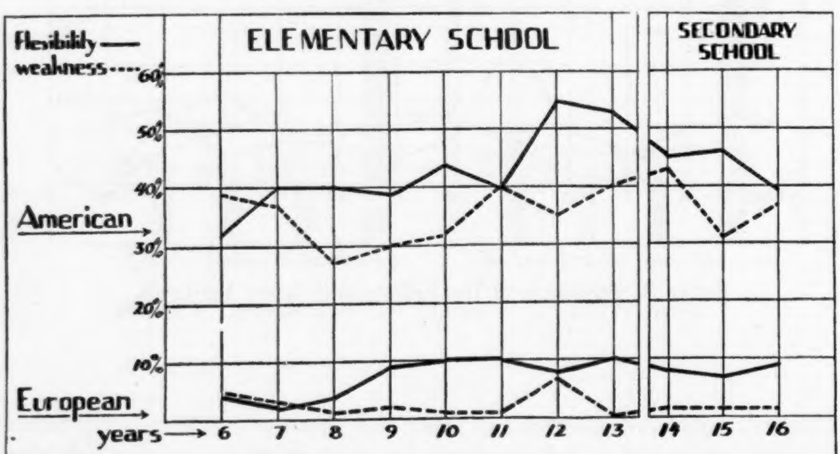


FIGURE IV. Flexibility Failures and Weakness Failures at Different Age Levels (Comparison of American versus Combined European Figures).

These figures are preliminary ones. We have started investigation of rural areas where the preliminary reports yield better results (32%) than those found in suburban and urban areas. Studies in private schools, where more time is given to physical education, also give better results.

At present we are in the process of evaluating tests given to children diagnosed as emotionally unstable. We are trying to determine whether a relationship exists between their emotional condition and their physical fitness levels.

Conclusions

We have the impression that insufficient exercise may cause the dropping of muscular fitness levels below the minimum necessary for daily living. The same lack of exercise may cause inadequate outlet for nervous tension.

Lack of sufficient exercise, therefore, constitutes a serious deficiency comparable with vitamin deficiency. Prevention of this deficiency is an urgent need.

Further research will be necessary to complete and broaden our preliminary survey, and to show the geographical incidence of under-exercise in this country. We may well find large sections where excellent conditions of muscular fitness prevail.

Some form of muscular fitness tests should be a part of every school health examination.[1]

Our physical education is in definite need of expansion so that there can be active and total participation not only in high schools, but even more important, in elementary school and pre-school groups.

REFERENCES

1. Cannon, Walter B. *The Wisdom of the Body.* New York: W. W. Norton & Co., Inc. 1939. 333 pp.
2. Gaston, Sawnie R. *Preliminary Report of a Group Study of the Painful Back.* New York: Columbia Presbyterian Medical Center.
3. ———. "The Low Back Syndrome," *Surgical Clinics of North America,* 31:2 (April 1951).
4. Gurewitsch, A. D., and Margaret A. O'Neill. "Flexibility of Healthy Children," *Archives of Physical Therapy,* April 1944.
5. Kendall, Henry A. and Florence P. "Normal Flexibility According to Age Groups," *The Journal of Bone and Joint Surgery,* 30A, pp. 690-694 (July 1949).
6. Kraus, Hans. "Diagnosis and Treatment of Low Back Pain," *GP,* V:4 (April 1952).
7. ———. "Prevention of Low Back Pain," *Journal of the Association for Physical and Mental Rehabilitation,* 6:1 (Sept.-Oct. 1952).
8. ———. "The Role of Therapeutic Exercises in the Treatment of Low Back Pain," *New York State Journal of Medicine* 49:13 (July 1, 1949).
9. ———. *Principles and Practice of Therapeutic Exercises,* Springfield, Ill.: Charles C. Thomas, 1949. 309 pp.

---

[1]The tests presented in this paper determine minimum fitness levels (6) and therefore constitute a good minimum requirement.

10. ——— and RUTH P. HIRSCHLAND. "Muscular Fitness and Orthopedic Disability." (Paper delivered before medical society of the State of New York, May 30, 1953, Buffalo, N. Y.—To be published in the *New York State Journal of Medicine.*)
11. ——— and RUTH P. HIRSCHLAND. "Muscular Fitness and Health," *Journal of the American Association for Health, Physical Education and Recreation,* 24:10:17 (Dec. 1953).
12. ——— and E. EISENMENGER-WEBER. "Evaluation of Posture Based on Structural and Functional Measurements," *Physio-therapy Rev.,* 25:6 (Nov.-Dec. 1945).
13. ——— and E. EISENMENGER-WEBER. "Fundamental Consideration of Posture Exercises," *Physio-therapy Rev.,* 27:6 (Nov.-Dec. 1947).
14. ——— and S. EISENMENGER-WEBER. "Quantitative Tabulation of Posture Evaluation," *Physio-therapy Rev.,* 26:5 (Sept.-Oct., 1946).
15. LORENZ, THOMAS H. and MARC J. MUSSER. "Life Stress and Painful Stiff Shoulders," *Modern Medicine,* March 1, 1953. (*From Ann. Int. Med.,* 37:1232-1244, 1952.)
16. SELYE, HANS. *The Physiology and Pathology of Exposure to Stress.* Montreal (Canada): Medical Publishers, 1950.
17. STIMSON, BARBARA. "The Low Back Problem," *Psychosomatic Medicine,* 9:3 (May-June 1947).
18. ———. "Backache," *American Journal of Nursing,* 51:11, (Nov. 1951).
19. ———. Personal Communications.
20. STINCHFIELD, FRANK. Personal Communications.
21. ——— and WILLIAM SINTON. "Criteria for Spinal Fusion," *Arch. Surgery,* Oct. 1952.

(Submitted 7-29-53)

Volleyball Skills of Junior High School Students as a Function of Physical Size and Maturity

NANCY A. LAMP

Garfield Junior High School
Berkeley, California

Abstract

This study investigated the volleyball playing ability of 806 junior high school students in relation to various physiological and growth factors. Statistical analysis of the volleyball tests showed them to be objective, reliable, and valid measures of playing ability. Positive correlations were found between volleyball playing ability (of both boys and girls) and the factors: age, height, weight, and strength. There was found to be a decided relationship between pubescent status and volleyball performance for both boys and girls. Pubescents and post-pubescents made higher scores at each age level than did pre-pubescents. A description of the volleyball tests is appended.

IT WAS THE purpose of this study[1] to investigate the relationship between (A) the factors of chronological age, physiological age, height, weight, and grip strength and (B) the volleyball playing ability of junior high school students, as determined by a battery of tests of fundamental volleyball skills.

It seems reasonable to believe that chronological age is related to playing ability, since at the junior high school period, the majority of physical skills do improve with age. However, physiological maturity is also an important factor in motor performance in adolescence and may be of greater significance than chronological age. Height may be of especial importance in volleyball performance in relation to reaching for and controlling balls. Weight is quite often a component part of an individual's strength and may play a part in the power of the return of the ball and of the serve. Strength, particularly hand and forearm strength, may be important in volleyball, perhaps more so than in other team games since the ball must be kept in the air by volleying the ball with the hands.

Review of the Literature

VOLLEYBALL

In an effort to determine the fundamental skills of volleyball, various articles and books about volleyball were studied and the frequency with which each different skill was suggested was tabulated. This tally of the fundamental skills was done in order to establish the terminology used most fre-

[1]This study was made in partial fulfillment of the requirements for the degree of Master of Arts in Physical Education, University of California, Berkeley, California, 1950.

quently by the various authors. The skills so named were the "serve", the "volley" (sometimes called the overhand pass), the "spike", and the "underhand pass" (occasionally called the "back relay," "net pass," or "underhand volley").

The survey reveals only four studies of test batteries which were developed statistically: Brady (6), Bassett, Glassow, and Locke (2), French and Cooper (14), and Russell and Lange (23).

Upon investigation of these studies, one is impressed by the fact that a "volleying" test has been included in each study, and that all but Brady included a serve test. The volley tests give consistently high coefficients of reliability and validity.

PHYSICAL SIZE AND MATURITY

Much of the existing information on human growth is based on "cross-sectional" data obtained on large numbers of individuals of varying ages, with resulting averages of height, weight, and pubescent status for boys and girls. A few longitudinal studies of boys and girls and their development have been published.

C. W. Crampton (7) was a pioneer in noting that the occurrence of pubescence did not always coincide with a specific chronological age. His study was published in 1908, and the criteria he used to determine pubescent age in boys have been used in many studies.

Studies on the relationship of maturation to other growth variables (strength, motor ability, body build, rate of growth) have been made by many investigators, among them: Atkinson (1), Bayley (3, 4), Bliss (5), Delaney (8), Dimock (9, 10), Espenschade (11, 12), Flory (13), Greulich (15), Jones (17, 18), Mullen (20), Nevers (21), Shuttleworth (24), and Wilson (25). Space does not permit a review of their findings here.

Procedure

In a pilot study of a battery of five volleyball skill tests administered to junior high school students, the test results were correlated with a composite rating of playing ability by two judges (physical education teachers of similar experience) as a criterion, and the sub-test scores were correlated with each other. The judges rated players independently. The tests which correlated highest with the criterion and lowest with the other sub-tests were retained. The method of scoring each test was revised on the basis of results in the pilot study so that there would be a wide range of scores, and the tests were made sufficiently difficult in order to avoid "perfect" scores. Four tests were retained (see Appendix, page 197).

MEASUREMENTS AND RATINGS

The grip-strengths of the right and left hands were determined for each student in the physical education classes. The elliptical hand-dynamometer, scaled in kilograms, was used for the test. The grip-strengths of the right and left hands were added to give the scores used for correlations with the volleyball test scores in this study.

The criteria recommended by Crampton (7) were used by the two boys physical education instructors in order to determine the pubescent status of each boy who took the volleyball tests. Three stages of pubescence are defined (termed P_1, P_2, and P_3 in this study): pre-pubescents, those boys who have not yet manifested any of the secondary sex characteristics; pubescents, those boys who have developed some hair in the pubic region; and post-pubescents, those boys who have kinky, pigmented hair in the pubic region.

In order to have three classification groups of girls comparable to those of boys, the criteria recommended by Delaney (8) were used. On the basis of these criteria, girls were grouped into three phases of pubescence (termed P_1, P_2, and P_3): pre-pubescents, those girls who have not yet manifested any of the secondary sex characteristics (who still have slender, boyish figures); pubescents, those girls whose breasts have begun to develop and whose pelvic girdles have broadened but who have not yet reached menarche; post-pubescents, those girls who have passed menarche.

These rather subjective judgments were made by two instructors who made individual ratings of the girls and then compared their ratings. Differences in decisions on P_1 and P_2 girls were rechecked and an agreement reached through discussion. A further check on the reliability of this subjective judgment was made six months later by rechecking the groupings made during volleyball season. Many of the P_2 girls had by then reached menarche, many of the P_1 girls were still pre-pubescent in the development of their figures, thus verifying the decisions reached earlier.

The chronological age, the height, and the weight of each student were recorded during the first two weeks of the study.

Subjects

The subjects used for this study were the students enrolled in the physical education classes at Garfield Junior High School in Berkeley, California, during the fall term of 1949. All students were required to take the volleyball practice and tests. Complete data were obtained on 429 girls and 377 boys in the seventh, eighth, and ninth grades. The age range for girls was from 11.3 to 15.0 years; for boys, from 11.0 to 16.5 years.

Administration of Tests

The two women instructors and one of the men instructors administered the tests to all of the students.

Discussion of Results

Objectivity of Tests

The objectivity of the tests was determined by having one class of girls (48 in number) tested by one teacher, then retested on the same day by a second teacher. The correlation coefficients of the two scores made by each girl on the tests are reported in Table 1.

The serve test and the volleying test do not require the assistance of another person in executing the test and are more objective in this respect than are the set-up and net-pass tests. In order to minimize the subjective element

involved because the tests require another individual to toss the ball to the subject, each of these two tests was administered to every student by the same instructor, except as noted in the special study groups.

Reliability of Tests

The reliability of the tests was determined by testing each girl in two other selected classes (90 in number) twice during the same physical education period, allowing an interval of at least ten minutes between test trials to prevent fatigue. The correlation coefficients of the two scores made by each girl on each test are lower than those obtained as measures of objectivity (see Table 1). This might be explained, in part, because the girls in this group were older and more familiar with volleyball techniques than were the girls in the first control group. If the older girls made a high score on the first trial, they did not try so hard to make a good score on the second test; or if they made a particularly poor score on the first trial, they were intent on making a high score on the retest.

However, the reliability coefficients in this study compare very favorably with the average of .660 obtained by French and Cooper (14) for their four-test battery of volleyball skill tests.

TABLE 1
Coefficients of Correlation for Determining:

	Reliability	Objectivity
Number of girls	90	48
Serve	.603	.750
Set-Up	.602	.721
Net-Pass	.561	.752
Volley	.644	.718

Validity of Tests

The validity of this battery of tests is indicated in two ways. The correlation of each test with the total score on the battery can be used with some reservation, if one allows for the fact that each test comprises part of the total score. The correlation coefficients of the girls' scores are higher than those of the boys' scores and indeed are extremely high (see Table 2). It would appear that any one of the tests for girls would be as effective as the entire battery in the measurement of volleyball ability. It is evident from the lower scores between tests and total score obtained from the boys' data that no one test can predict the total accurately. Evidently the skills are more specific in boys than in girls, as indicated by lower interrelationship among the sub-tests.

A second method of establishing the validity of the tests is that of comparing the test scores which were made by the girls who were selected to play on the teams representing the school in the inter-school and intra-mural tournaments. In each of the upper grades, the team girls made scores approximately one standard deviation above the mean. The mean for standard scores is 50, and the mean for the total score on four tests is 200. The stand-

TABLE 2
Correlations and Inter-correlations

BOYS	Number: 377							
Test	Serve	Set-Up	Net-Pass	Volley	Age	Ht.	Wt.	Str.
Serve	-----	-----	-----	-----	.570	.490	.367	.457
Set-Up	.420	-----	-----	-----	.343	.276	.312	.365
Net-Pass	.455	.324	-----	-----	.385	.416	.271	.366
Volley	.525	.496	.436	-----	.265	.643	.314	.400
Total	.787	.736	.710	.805	.514	.618	.395	.546

GIRLS	Number: 429							
Test	Serve	Set-Up	Net-Pass	Volley	Age	Ht.	Wt.	Str.
Serve	-----	-----	-----	-----	.705	.343	.539	.402
Set-Up	.548	-----	-----	-----	.942	.391	.506	.324
Net-Pass	.594	.539	-----	-----	.712	.440	.539	.372
Volley	.602	.603	.592	-----	.963	.470	.520	.358
Total	.974	.973	.957	.985	.877	.530	.749	.563

ard deviation of each test was ± 10; the standard deviation of the total was ± 31.03. An example from the table might better show how this indicates the validity of tests. (The mean for each test is 50; the mean for the total is 200).

L-9 Girl	*Serve*	*Set-Up*	*Net-Pass*	*Volley*	*Total*
A	63	69	64	62	258

Many of the seventh-grade-team girls made scores well above the mean for the entire school, even though these girls were younger and less experienced. These facts show that the better players make better scores on the tests, indicating that the tests are valid measures of volleyball playing ability.

Sex Differences

The mean scores and the standard deviations for the volleyball tests indicate that boys and girls make similar scores. There is, in fact, no significant difference in scores for any test, as indicated by very small critical ratios obtained in comparing the test scores of boys with the test scores of girls. (See Table 3).

TABLE 3
Critical Ratios of Test Scores
(Boys vs. Girls)

Serve	.546
Set-Up	.347
Net-Pass	.171
Volley	1.292
Total	.758

In order to be significant at the .01 level, the critical ratio for this group would have to be 2.59; at the .05 level, 2.33.

GROWTH FACTORS (See Table 2)

1. The correlations between age and the various tests were very high for girls. This contrasts with the rather low correlations for boys. This might be explained in part by the requirement made by the girls' physical education instructors that every girl practice and play volleyball four days a week for eight weeks of each year. The boys are not required to play volleyball, but are given a choice of activities. Those planning to try out for the teams are urged to play volleyball during their physical education classes. For the purpose of this study, the boys' instructors arranged for the boys to practice and to take the tests for scores. Therefore, there were many boys in the upper grades who had had no more practice than had the younger students. This relative lack of practice for the boys as compared with that for the girls might explain the higher correlations for girls between test scores and total scores and the higher correlations between the tests and the factor of age, which in each case was markedly higher for the girls than for the boys.

2. The factor of height for both boys and girls shows more relationship to the volleying test than to any of the other tests, substantiating an hypothesis that was mentioned in the introduction.

The set-up test had the lowest correlation with height for the boys, and next lowest for the girls. This test demands accuracy of placement rather than height or strength.

The serve test had the lowest correlation with height of any of the tests for girls, but the second highest for boys. Neither coefficient is high, however. Strength appears to be slightly more important than height for the serve test for girls. For boys, these two factors are approximately equal in importance.

3. For the boys, the correlations between weight and the various tests were the lowest of the factors studied. For the girls, weight is evidently more important than height or strength. Strength has been found to vary with weight (19) and it should be noted that for both boys and girls the highest correlations between both of these factors and the tests are with the serve, followed, in order, by the volley, net-pass and the set-up. The total score for the girls correlated highly with weight but that correlation for boys is low.

4. The relationship between strength and the various tests was consistent, although low for both the boys and the girls. As mentioned above, strength and weight showed similar relationships with the tests for both boys and girls. The serve test correlated more highly with strength than did the other tests, which is not surprising since, by analysis, strength is essential in getting the ball over the net into the back of the opponents' court for a higher score on the test. The total score had the highest correlation with strength and this fact indicates that strength is important for volleyball playing ability.

PUBESCENCE

The percentages within each pubescent group for each half-year interval and for each grade are given in Table 4.

TABLE 4

Percentage of Each Pubescent Group for Each Half-Year Interval

Age	*Boys* Number: 377			*Girls* Number: 429		
	P_1	P_2	P_3	P_1	P_2	P_3
11.0–11.5	.8	.5	----	.5	----	.2
11.6–11.9	1.3	1.3	----	1.2	.9	.7
12.0–12.5	7.2	8.5	.3	5.6	4.0	5.8
12.6–12.9	3.4	9.0	5.0	3.7	3.0	8.6
13.0–13.5	1.6	8.8	5.6	2.8	3.5	13.1
13.6–13.9	.5	5.0	8.2	.5	.7	13.1
14.0–14.5	----	6.1	8.5	.7	.2	13.7
14.6–14.9	.3	1.6	8.8	.5	.2	10.3
15.0–15.5	----	1.6	4.0	----	----	3.0
15.6–15.9	----	----	1.3	----	----	.5
16.0–16.5	----	----	.8	----	----	----
Total	15.1	42.4	42.5	15.5	15.5	69.0

Percentage of Each Pubescent Group for Each Grade in School

Grade	*Boys* Number: 377			*Girls* Number: 429		
	P_1	P_2	P_3	P_1	P_2	P_3
L7	11.7	11.7	.5	8.6	5.5	7.0
H7	----	7.2	5.0	1.7	2.6	7.0
L8	3.2	12.5	3.0	3.8	3.9	16.6
H8	----	3.2	11.3	.7	2.3	8.6
L9	.2	4.6	16.2	.7	.7	21.0
H9	----	3.2	6.5	----	.5	8.8
Total	15.1	42.4	42.5	15.5	15.5	69.0

The composite test score for each student was tabulated on a two-way chart with pubescent status (P_1, P_2, P_3) as one variable and age as the second variable. The mean scores of each group were charged (Figures I and II) for purposes of comparison.

It may be seen (Figure 1) that the mean test scores for the post-pubescent boys are higher at each age level than are the mean scores of the pubescent boys, and these latter are, in turn, higher at each age level than those of the pre-pubescent boys. It is interesting to note that the pre-pubescent boys show little increase in performance with age. Indeed there is an indication of less in score with age, in this group. Both of the other groups increase with age, but the post-pubescent group shows an especially marked superiority in the 14th year.

The figure for girls (Figure II) indicates that the pubescent group achieved higher mean scores than did either the pre-pubescent or post-pubescent girls in the age level 12 to 13.5, after which age the more mature girls assumed and maintained a superiority over the less mature girls. The mean scores of

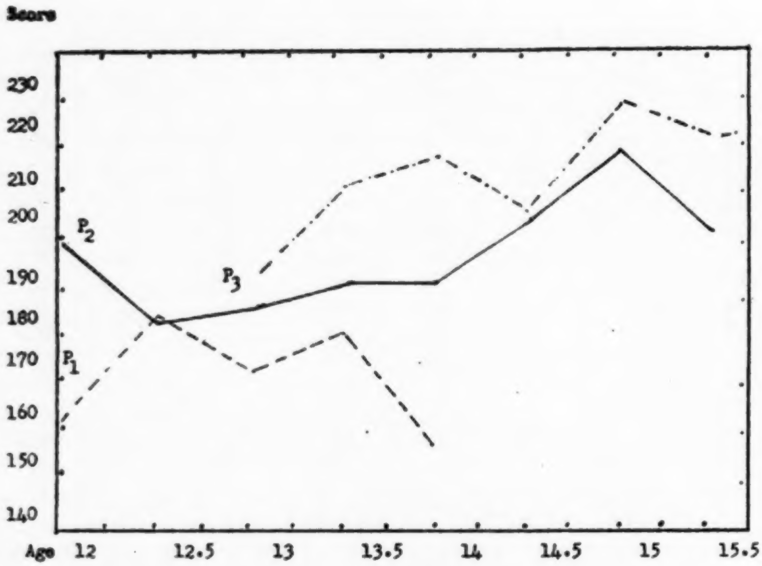


FIGURE I. Volleyball Scores in Relation to Pubescence (Boys)

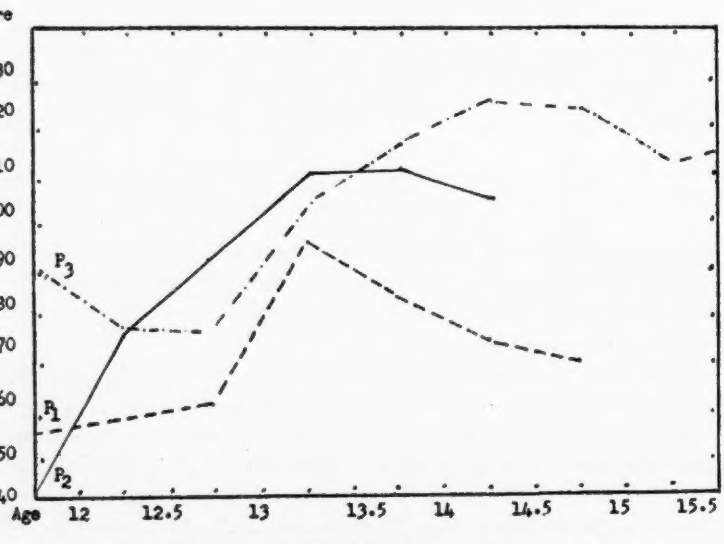


FIGURE II. Volleyball Scores in Relation to Pubescence (Girls)

the pre-pubescent girls were at all ages less than the mean scores of the pubescent and post-pubescent girls. All groups show an early increase with age and in all groups the maximum increase appears to come between 12.75 and 13.25 years. After this period, the pre-pubescents show a leveling off or

even a decline in scores. Peak scores for both pubescent and post-pubescents appear to come in the 15th year, followed by a leveling off or decline. These results substantiate evidence published by Espenschade (12) that girls improve in ability up to the age of 14, then tend to decline. Nielson and Cozens (22) also showed that girls tend to plateau, then double-back or decline in ability after the age of 14.

Conclusions

1. The volleyball tests are reasonably objective as evidenced by the correlation coefficients of the order of .7 (see Table 1).
2. The tests are as reliable as those used by other investigators of volleyball skills. Correlation coefficients of the order of .6 were obtained (see Table 1).
3. The tests are valid as indicated by the correlation coefficients with total score of the order of .9 for the girls and .7 for the boys. The scores made by the "team" girls were approximately one standard deviation above the mean score, which indicates that, for girls, the tests do discriminate among players of varying ability.
4. There is no significant difference between boys and girls in their ability at this age to perform the skills of volleyball, as evidenced by the very small critical ratios comparing the scores made by the boys with those made by the girls.
5. Age and weight are more closely related for girls than for boys in performance in volleyball skills.
6. Height is more important than the other growth factors for boys in relation to volleyball skill tests.
7. For both boys and girls there is a slight positive relationship between strength and volleyball playing ability.
8. A comparison of scores and pubescent status indicates that there is a decided relationship between these factors for junior high school boys. The more mature boy at each chronological age scores higher than the less mature boy.

For the girls, all pubescent groups show an early increase in performance with age and in all groups the maximum increase appears to come between 12.75 and 13.25 years. Peak scores for the pubescent and post-pubescent groups appear to come in the 14th year, followed by a decline or leveling off of scores.

APPENDIX

VOLLEYBALL TESTS

The Serve Test

Equipment needed: A number of properly inflated volleyballs—at least one ball for every ten students; standard volleyball courts, 30′ x 60′ with taut nets 7′6″ high at the center.

Court markings: A chalk line parallel to and 15′ from the net, dividing the court into two equal halves.

(This test is described for right-handed players. Reverse co-ordinations for a left-handed player.) The subject stands outside the right rear corner of the court. He stands with feet together, facing the net, the ball in his left hand, his right hand under the left. He steps back with the right foot, bending forward at the waist, the ball lowered to knee height, the right hand raised backward. As he serves, the weight is transferred forward, the body is lifted as the right hand replaces the left under the ball and, with a slap, the ball is served over the net. The ball can be directed by turning the body so that the follow-through of the right hand is in the direction desired for the placement of the ball.

Specific requirements for scoring on each trial:

1. The serving hand must be open, so that the fingers can help direct the ball.
2. The ball must *not* be tossed up by the left hand before being served.
3. The ball must not touch the net. Trial scores 0 in this case.
4. The server must not step on or over the back line during the serve.
5. Balls landing in the rear section of the opponents' court or on the chalk line score two points.
6. Balls landing in the front section of the opposite court score one point.
7. Ten trials shall be allowed and scores on each trial recorded.

The Set-Up Test

Equipment needed: A number of properly inflated volleyballs—at least one ball for every ten students; standard volleyball courts, 30′ x 60′ with taut nets 7′6″ high at the center; two jumping standards, with a cross-bar or rope placed at the height of 7′6″.

Court markings: Chalk lines 10′ from and parallel to the side lines, dividing the court into three equal strips. A chalk line 6′ from and parallel to the net. Each of the lengthwise strips may be used for the test, hence three students could be tested at one time on one court. The jumping standards are placed on the intersections of the chalk lines, so that the rope or cross-bar is 6′ from and parallel to the net, and the testing area is 10′ wide.

The subject stands in the rear half of one of the strips, facing the net, the cross-bar and front half of the strip between him and the net. The tester stands in the opponents' court, near the net *and* the rope or cross-bar to the subject, who must set-up the ball in such a manner that the ball goes over the rope or cross-bar and lands in the 6′ x 10′ area between the rope and net.

Specific requirements for scoring on each trial:

1. The ball must be hit with both hands.
2. The thumbs must be adjacent when volleying the ball. An underhand pass does not score.
3. The ball cannot be held, but must be clearly batted.
4. Balls landing in the designated area count two points.
5. Balls touching the rope or cross-bar or any of the lines bounding the scoring area count one point.
6. Ten trials shall be allowed and scores for each trial recorded.

The Net-Pass Test

Equipment needed: A number of properly inflated volleyballs—at least one ball for every ten students; standard volleyball courts, 30′ x 60′ with taut nets 7′6″ high at the center; two jumping standards, with a cross-bar or rope placed at the height of 7′6″.

Court markings: Chalk lines 10′ from and parallel to the side lines, dividing the court into three equal strips. A chalk line 6′ from and parallel to the net. Each of the lengthwise strips may be used for the test, hence three students could be tested at one time on one court. The jumping standards are placed on the intersections of the chalk lines, so that the rope or cross-bar is 6′ from and parallel to the net. The opponents' court is divided crosswise by a chalk line 15′ from and parallel to the net.

The subject stands with his back to the net in the front half of one of the three 10' wide strips (between the cross-bar and the net). The tester stands in the rear half of the same strip, facing the subject. The tester tosses the ball over the rope or cross-bar in such a manner that the ball, if allowed to drop, would land in the middle of the 6' x 10' rectangle in which the subject is standing. As the ball is tossed over the rope, the subject brings his hands together, palms upward and little fingers adjacent, and bats the ball back over his own head and over the net, so that the ball will land in the opponents' court.

Specific requirements for scoring on each trial:

1. The ball must be clearly batted, not thrown or held.
2. Both hands must be used in batting the ball.
3. Balls landing on the chalk line or in the rear half of the opponents' court count two points.
4. Balls landing in the front half of the opponents' court count one point.
5. Balls touching the net and going into the opponents' court count according to the area in which they land.
6. Ten trials shall be allowed and score for each trial recorded.

The Volleying Test

Equipment needed: A number of properly inflated volleyballs—at least one ball for every ten students; a tin strip nailed to the wall at a height of 7'6" or a chalk line marked on the wall at this height; a stop-watch and a whistle.

Other markings required: A chalk line on the floor 3' from and paralled to the wall. The line should be the length of the 7'6" strip or line on the wall (at least 10').

The subject stands behind the line 3' from and facing the wall, holding the ball in one or both hands. On a signal from the tester, the subject tosses the ball against the wall above the 7'6" line. As the ball rebounds from the wall, the subject tries to volley the ball repeatedly above the line as often as possible during the 15 seconds allowed.

Specific requirements for the test:

1. Each time the ball is batted against the wall above the 7'6" line shall be counted.
2. The ball should be batted with two hands, but one-handed volleys shall count.
3. If the ball gets out of control, it must be recovered and brought back to the 3' line to be started again.
4. The subject may move as close to the wall as he feels necessary in order to volley the ball against the wall successfully.
5. The score for the test shall be the number of times the ball is volleyed *above* the 7'6" line during the 15 seconds. (Balls landing on the line do not score.)

REFERENCES

1. Atkinson, R. K., A Study of Athletic Ability of High School Girls, *American Physical Ed. Rev.*, **30**: 389-399 (Sept. 1925).
2. Bassett, G., R. Glassow, and M. Locke, Studies in Testing Volleyball Skills, *Research Quarterly*, **8**: 60-72 (Dec. 1937).
3. Bayley, N., Skeletal Maturing in Adolescents as a Basis of Determining Percentage of Completed Growth, *Child Development*, **14**: 1-46 (March 1943).
4. ———, Size and Body Build of Adolescence in Relation to Rate of Skeletal Maturing, *Child Development*, **14**: 47-90 (March 1943).
5. Bliss, J. S., A Study of Progression Based on Age, Sex, and Individual Differences in Strength and Skill, *American Physical Ed. Rev.*, **32**: 11-21 (Jan. 1927); 85-89 (Feb. 1927).
6. Brady, G. F., Preliminary Investigation of Volleyball Playing, *Research Quarterly*, **16**: 14-17 (March 1945).
7. Crampton, C. W., Physiological Age—A Fundamental Principle, *American Physical Ed. Rev.*, **13**: 141-154 (March 1908).

8. Delaney, M., Age, Height, Weight and Pubescent Standards for Athletic Handicapping of Girls, *American Physical Ed. Rev.*, **33:** 507 (Oct. 1928).
9. Dimock, H., Research in Adolescence, *Child Development*, **6:** 176-185 (Sept. 1935).
10. ———, Strength and Motor Ability in Relation to Pubescence, *Child Development*, **6:** 186-195 (Sept. 1935).
11. Espenschade, A., *Motor Performance in Adolescence.* Mono. of Society for Research in Child Development, **5:** 1-126 (1940).
12. ———, Development of Motor Coordination in Boys and Girls, *Research Quarterly*, **18:** 30-43 (March 1947).
13. Flory, C. D., Predicting Puberty, *Child Development*, **6:** 1-6 (March 1935).
14. French, E. L., and B. I. Cooper, Achievement Tests in Volleyball for High School Girls, *Research Quarterly*, **8:** 150-157 (May 1937).
15. Greulich, W. W., Physiological Changes in Adolescence, *Adolescence*, 43rd Yearbook, National Society for Study of Education, Chicago, Illinois: University of Chicago Press, 8-32 (1944).
16. Jones, H. E., Development of Physical Abilities, *Adolescence*, 43rd Yearbook, National Society for Study of Education, Chicago, Illinois: University of Chicago Press, 100-121 (1944).
17. ———, The Sexual Maturing of Girls as Related to Growth in Strength, *Research Quarterly*, **18:** 135-143 (May 1947).
18. ———, *Motor Performance and Growth.* Berkeley, California: University of California Press, 1949.
19. McCloy, C. H., The Apparent Importance of Arm Strength in Athletics, *Research Quarterly*, **6:** 61-64 (May 1935).
20. Mullen, F., Factors in the Growth of Girls, *Child Development*, **12:** 27-42 (March 1940).
21. Nevers, J. E., The Effects of Physiological Age on Motor Achievement, *Research Quarterly*, **19:** 103-10 (May 1948).
22. Neilson, N., and F. W. Cozens, *Achievement Scales in Physical Education*, A. S. Barnes Co., 3-19, 1934.
23. Russell, N., and D. Lange, Achievement Tests in Volleyball for Junior High School Girls, *Research Quarterly*, **11:** 33-41 (Dec. 1940).
24. Shuttleworth, F. K., *Sexual Maturity and the Physical Growth of Girls, 6 to 19.* Mono. of Society for Research in Child Development, 11 (1937).
25. Wilson, M., A Study of Arm and Shoulder Girdle Strength of College Women in Selected Tests, *Research Quarterly*, **15:** 258-267 (Oct. 1944).

(Submitted 6-13-52)

On the Use of the Mean and Median in Stop Watch Timing[1]

HENRY J. MONTOYE
Michigan State College
East Lansing, Michigan

BUFORD BECK
University of Illinois
Urbana, Illinois

Abstract

Present athletic practices in stop-watch timing require the median of three watch times to be declared "official" time. Theoretical evidence is available that the standard error of the median is larger than the standard error of the mean when sampling is from a single, normal population. The present investigation was designed to determine if, under actual race conditions, the same relationship between the mean and median time exists. In the two independent experiments described the standard error of the median was observed to be larger than that of the mean.

IN AN INTERESTING article by Henry (4), the results of an experiment in stop watch timing were reported. In this investigation, movement of the sweep-second hand of an electric clock was employed as a criterion of true time while a series of 150 consecutive 10-second measurements was made by an observer with a stop watch. The stop watch times thus obtained were arranged as serial triads. True time was estimated more accurately from the arithmetic means of a triad than from the middlemost score, mid-range, or by any of the several methods of discarding discrepant values. When the data were arranged in groups of five, the arithmetic mean again yielded more accurate estimates of true time than any of the other methods.

The standard deviation of the errors (i.e., differences between true time and watch time) was 0.084 seconds using the means of the triads. This may be considered the average standard error of the mean with $N=3$. The middlemost (median) scores of the triads resulted in a σ of 0.112 seconds or 1.33 σ_M. The ratio of the variances, 1.78, is significant at the 5-per-cent level.

The Amateur Athletic Union, Intercollegiate Amateur Athletic Association of America, National Collegiate Athletic Association, and National Section for Girls and Women's Sports require that the median of three times be declared official time in races of various kinds. One may be tempted to infer from the work of Henry that current practices in the handling of stop watch times do not produce most accurate results. A moment of reflection, however, reveals important differences between this laboratory experiment

[1]A portion of this study appeared in a Master's thesis by Buford Beck entitled *An Investigation of the Validity of Certain Practices in Stop Watch Timing,* Michigan State College, 1950.

and the timing of a typical track event, for example. While the results of Henry's research no doubt apply when repeated measurements are taken by one observer (blood counts, pulse wave measurements, etc.), their direct application in a race such as a hundred-yard dash needs verification.

In Henry's study, since the stimulus for starting and stopping the watch was identical, a serious bias in timing errors is not expected. (There may be a small bias, however, due to imperfection in the mechanism of the stop watch.) This is not the situation when athletic events are timed, especially when the timers are not among the best. Previous studies (1, 2, 3, 5, 10) indicate that, in general, an appreciable bias does exist. This bias appears to be due to a combination of factors, most important of which is the delay in starting the watch at the beginning of the race. There is no compensatory delay at the end of the race since most timers apparently anticipate the finish. The result is to underestimate true time.

In the laboratory experiment, inasmuch as only a single observer was employed, sampling was from one population. In a track race, more than one population may be involved. It is logical that these factors may produce results different from those secured by Henry. It was the purpose of this study to observe the effects of these factors on the mean and median as estimates of true time in a typical track race.

Statistical Theory

The standard error of the median (σ_{Md}) is given as 1.2533 σ_M in most statistical texts. This is the limiting value, however, and applies only when sampling is from a *single*, normal population. When *small* samples are taken from a single, normal population, the sampling distribution of the median is also known (6, 7, 8). Several formulae have been developed for estimating the σ_{Md}. They give approximately the same results. Hojo's (6) are:

$$\sigma_{Md} = (\sigma/\sqrt{N})\ (1.2533 - .2653/N - 0699/N^2 + .0822/N^3) \text{ when N is odd}$$
$$\sigma_{Md} = (\sigma/\sqrt{N})\ (1.2533 - .8261/N + .7826/N^2 - .3478/N^3 + .1304/N^4) \text{ when N is even}$$

For samples of three, the σ_{Md} is 1.1602 σ_M. Samples of 5 result in a σ_{Md} of 1.1976 σ_M. In order to verify the theoretical equations, Hojo, using Tippetts numbers, selected 1,000 samples of 3. The standard deviation of the medians of these samples was 1.1321 σ_M which is quite close to the theoretical value of 1.1602 σ_M.

When the parent population is not normal, somewhat different results may be expected. Pearson and Adyanthaya (7) determined the σ_{Md} for small samples in terms of the σ_M when $\beta_2 = 1.80$ (i.e., when the distribution is rectangular). For samples of 3 taken from this population, the $\sigma_{Md} = 1.342\ \sigma_M$ and for samples of 5 the $\sigma_{Md} = 1.464\ \sigma_M$. Using Tippetts numbers, these authors also extracted 500 samples of 5 from each of three populations having β_2 values of 2.5, 4.122, 7.069. The standard errors of the medians were

1.319 σ_M, 1.145 σ_M, and 1.049 σ_M respectively. Apparently, the advantage of the mean over the median becomes less as the distribution becomes more leptokurtic.

The experiment of Henry's probably approximates the conditions for which these statistical formulae have been developed. The data were secured from a single population which very likely approached normality. The present experiment differed from Henry's and conceivably from the conditions for which the theoretical formulae apply.

Experiment 1

METHOD

An outdoor, 50-yard straightaway lane was marked off with a starting line, starting blocks, and an official finish line. Each runner was started with an open-muzzled, .22 calibre starting gun and was timed by five timers and an electric timing device. "True time" as recorded by the electric timing device, and watch times were read and recorded by a sixth observer. The timers were not permitted to observe the times marked by the stop watches or the electric timer.

The electric timing device[2] constructed for this experiment consisted of a starting switch operated by the muzzle blast of the starting gun, an electric chronometer which was activated by the action of the starting switch, and a properly shaded photoelectric cell unit at the finish which stopped the chronometer when the beam was broken. Inasmuch as times marked by the electric timer were considered "true time," an investigation was made of the source and magnitude of the errors contained within this system. An analysis of the errors due to (1) changes in the power source, (2) activation of relay circuits, (3) velocity of the explosive waves, (4) sensitivity of the photoelectric cell, and (5) position of the yarn and the photoelectric cell at the finish resulted in an estimated error of .04 seconds, if all errors were operating at a maximum. The direction of the error was such as to cause the electric chronometer times to be greater (i.e., slower) than watch times.

The stop watches, recently calibrated, were those used in the AAU and NCAA sanctioned track and swimming events held at Michigan State College. They are, in all probability, typical of those used in the average college or high-school track meets.

The timers included the assistant track coach and assistant wrestling coach at Michigan State College, a high-school track coach, a graduate student, and one of the present writers, all of whom had had considerable experience in timing. The timers were located on a three-step judges stand on one side and a natural slope on the other side of the finish line.

---

[2]The authors wish to express their appreciation to Robert Clark, graduate student in electrical engineering at Michigan State College, who constructed the electric timer.

RESULTS

The timing errors for each of the five timers are summarized in Table 1. Application of Bartlett's test of homogeneity of variances (9, *p. 249*) resulted in a corrected chi square of 12.2 which is significant between the 1- and 2-per-cent levels. This indicates a significant difference among timers in the viability or spread of their errors.

TABLE 1
Mean, Variance, Standard Deviation and Range of Errors[1] *for the Five Timers*
(N = 101 for each timer)

Timer	Mean (*sec.*)	σ^2 (*sec.*)	σ (*sec.*)	Limits (*sec.*)
A	—.100	.0079	.089	—.3 to +.2
B	—.127	.0071	.084	—.3 to 0.0
C	—.141	.0120	.109	—.5 to +.1
D	—.073	.0078	.088	—.3 to +.1
E	—.045	.0116	.108	—.3 to +.2

[1]Watch time minus "true" time.

The critical ratio technique was applied in analyzing the differences in *mean* errors among the timers. For a difference to be significant, it must be 0.027 at the 5-per-cent level and 0.036 at the 1-per-cent level. The means of the errors of the five timers are therefore significantly different except possibly in the comparison between timers B and C. All of the means were significantly greater than zero at the 1-per-cent level. Hence, we may conclude (a) four and possibly five populations are being sampled and (b) there is a bias causing watch times to be shorter than true time. The direction of the bias is in agreement with most previous studies (2, 5). The bias probably includes the small error in the electric timer.

We do not have *a priori* knowledge of how the sampling of several populations affects the relationship between the standard deviation of the median and that of the mean when three or five timers are being used. Obviously, it will depend upon the particular timers employed. For example, if two timers were consistently close to true time and one consistently deviated to a considerable extent in one direction, the median would be a better estimate of true time. However, this is probably an unlikely occurrence.

In handling the data in the present experiment, ten combinations of the timers were formed. Each combination included 3 timers who might be considered "official" timers. Their times on each of the 101 trials were averaged by taking the median and the mean.[3] True time was then subtracted from each of the averages thus obtained. The result was a distribution of errors of the median and a distribution of errors of the mean for each of the ten combinations of timers. The results are presented in Tables 2 and 3.

[3]The averages were carried out only to the closest tenth of a second. Although some precision is lost in this manner, in most track situations with times measured in tenths, it serves no practical purpose to record the averages of three times in hundredths.

TABLE 2
Means, Variances, and Standard Deviations of Errors Using the Median and Mean as Estimates of True Time

Timer	Errors of Median Mean (sec.)	(sec.) σ^2	(sec.) σ	Errors of Mean Mean (sec.)	σ^2 (sec.)	σ (sec.)
ABC	—.113	.0062	.079	—.119	.0060	.077
ABD	—.097	.0055	.074	—.101	.0041	.064
ABE	—.089	.0064	.080	—.088	.0056	.075
ACD	—.090	.0067	.082	—.101	.0056	.075
ACE	—.086	.0068	.082	—.094	.0053	.073
ADE	—.057	.0066	.081	—.069	.0057	.075
BCD	—.102	.0065	.081	—.107	.0059	.077
BCE	—.098	.0069	.083	—.093	.0062	.079
BDE	—.079	.0064	.080	—.078	.0049	.070
CDE	—.073	.0073	.085	—.082	.0056	.075

TABLE 3
Comparison of the Mean and Median of Watch Times as Etimates of True Time

Timers	Difference in Mean Error (Md — M)	Variance Ratio (σ^2Md ÷ σ^2M)	Ratio of Standard Deviations (σMd ÷ σM)
ABC	.006	1.03	1.03
ABD	.004	1.34	1.16
ABE	—.001	1.14	1.07
ACD	.011	1.20	1.09
ACE	.008	1.28	1.12
ADE	.012	1.16	1.08
BCD	.005	1.10	1.05
BCE	—.005	1.11	1.05
BDE	—.001	1.31	1.14
CDE	.009	1.30	1.13
Weighted Mean	.0048	1.19	1.09

The variance ratios of Table 3 are obviously not distributed about a mean of 1.00 as would be the case if there were no difference in variance between the means and medians of three times. Three of the F-ratios have probabilities slightly over 0.05. When a single F-ratio is computed, its value has a probability of less than 0.01. Therefore, with these combinations of timers, the means of the three watch times produced a smaller variance or random error than when the median was so employed. The increase in σ using the medians was 9 per cent, which is slightly smaller than the 16 per cent expected when sampling is from a single, normal population.[4] This is closer to the theoretical value, however, than the 33 per cent secured by Henry (4), although the difference probably represents sampling variation.

[4]Rounding off the averages reduced the advantage of the mean. When these averages were carried out, the increase in σ of the medians as compared to the means was 13.5 per cent.

The variances of the averages were considerably smaller, in general, than the variances of the errors of the individual watch times. This, of course, is not surprising in view of the relationship between the standard deviation and standard error of the mean.

Identical statistical analysis was carried out for the averages of five watch times. The variances of the errors were 0.0053 and 0.0037 for the medians and means respectively. As anticipated, these values are appreciably smaller than the variances of the errors of the individual times and averages of three watch times. The ratio of the variances for medians and means (5 watches) is significant between the five and one per cent levels. The corresponding ratio of standard deviations of the errors is 1.198, an increase in size of the σ of the errors of the medians of 19.8 per cent as compared to the σ of the errors of the means. This is precisely the value expected when sampling is from a single, normal population.

Experiment 2

METHOD

Approximately two years after the first experiment was completed, an improved electronic timing unit was available. It was considered advisable to repeat the stop-watch study as a check on the original results. Modifications of the timing apparatus[5] included (a) greater sensitivity of the photoelectric cell, (b) activation of the chronometer by sound rather than air blast from starting gun, and (c) a change in position of the tape and photoelectric cell at the finish line. The computed maximum error of the new apparatus was less than 0.01 seconds.

In the second experiment, the five college students who operated the stop watches had had no previous experience in timing races. They were squad leaders selected from a physical education class who were given a few minutes of instruction and a brief practice period. The subjects who ran the 50-yard races were also college men. Two subjects ran at a time, but only the lead runner was timed by the chronometer and stop watches. A total of 65 races were timed.

The second experiment was identical with the first in all other respects.

RESULTS

The results were analyzed just as before and appear in the analogous Tables 4, 5 and 6. Bartlett's test was applied to the data in Table 4 resulting in a statistically significant (2% level) chi square of 11.45. This indicates that the variances (σ^2) of the timers' errors differed significantly.

There was also a significant difference in the mean errors of the timers at the 1-per-cent level in four of the ten comparisons. With each timer, the error was significantly different from 0, indicating the presence of a bias causing watch times to be shorter (faster) than true time. In the second experiment,

---

[5]Details of this electronic timer are described in a paper being prepared for publication.

therefore, a significant bias in watch times was present, as well as a difference among timers in the magnitude of this bias and the variance of the random errors. This is much the same state of affairs as existed in the experiment with experienced timers.

The variances of the errors were greater for the inexperienced timers of the second experiment. This is not surprising since the "experts" are probably trained to be more consistent in their attentiveness and reaction. The bias was generally less in the second experiment. This is explained by the improved timber which practically eliminated errors in the criterion of true time.

As in the first experiment, the means of the three times resulted in smaller random errors than did the medians. The mean increase in size of σ for the medians as compared to the means was approximately 17 per cent. This is again very close to the 16 per cent expected if sampling were from a single, normal population.

When the watch times of all five timers were averaged together, the variance of the errors of the means was 0.0063 and for the medians, 0.0084. Corresponding standard deviations were 0.0794 and 0.0916. The increase in σ for the medians as compared to the means was 16 per cent: the theoretical value for a single, normal population being approximately 20 per cent.

TABLE 4

Mean, Variance, Standard Deviation, and Range of Errors[1] For the Five Watches
(N = 65 for each watch)

Watch	Mean (sec.)	(sec.) σ	(sec.) σ^2	Limits (sec.)
A	−.074	.0113	.106	−.4 to +.2
B	−.048	.0181	.134	−.4 to +.2
C	−.134	.0212	.145	−.5 to +.2
D	−.042	.0119	.109	−.3 to +.3
E	−.046	.0209	.144	−.4 to +.3

1Watch time minus "true" time.

TABLE 5

Mean, Variance, and Standard Deviation of Errors Using the Median and Mean as Estimates of True Time

Timer	Errors of Median			Errors of Mean		
	Mean	σ^2	σ	Mean	σ^2	σ
ABC	−.068	.0136	.1064	−.083	.0091	.0954
ABD	−.057	.0095	.0976	−.045	.0080	.0894
ABE	−.051	.0126	.1124	−.058	.0066	.0812
ACD	−.069	.0080	.0892	−.078	.0032	.0568
ACE	−.080	.0102	.1010	−.086	.0083	.0909
ADE	−.055	.0077	.0878	−.051	.0080	.0896
BCD	−.062	.0104	.1017	−.066	.0074	.0860
BCE	−.058	.0141	.1187	−.080	.0102	.1009
BDE	−.040	.0104	.1018	−.046	.0081	.0900
CDE	−.069	.0092	.0958	−.077	.0070	.0836

TABLE 6

Comparison of the Mean and Median of Watch Times as Estimates of True Time

Timer	Difference in Mean Error (Md — M)	Variance Ratio (σ^2Md ÷ σ^2M)	Ratio of Standard Deviations (σMd ÷ σM)
ABC	.015	1.25	1.12
ABD	—.012	1.19	1.09
ABE	.007	1.91	1.38
ACD	.009	2.50	1.57
ACE	.006	1.23	1.11
ADE	—.004	0.96	0.98
BCD	.004	1.40	1.18
BCE	.002	1.38	1.18
BDE	.006	1.28	1.13
CDE	.008	1.31	1.14
Weighted Mean	.0061	1.393	1.172

Discussion

It is apparent that the reduction in error variance effected by employing the mean of three or five stop-watch times, rather than the median is approximately equal to that expected when sampling is from a single, normal population. The experimental results were surprisingly close to the theoretical values, despite significant differences in the individual bias and variance of the timers' errors. The two independent track-timing experiments yielded comparable results.

There was, however, one disturbing effect. When three timers were employed, the medians produced a statistically significant smaller bias than the means. In other words, the mean error was less for the medians. Both experiments resulted in significant critical ratios; namely, 3.36 and 2.46. This occurred despite the greater variance for the medians. However, the difference in mean bias for the medians and means was quite small (see Tables 3 and 6). When five watches were averaged, the mean bias for the medians and means was not significantly different.

An inspection of the distribution of medians and means showed a slight negative skewness for the medians which was present in both experiments. However, there was no statistical evidence that the distributions of the medians and means differed significantly. This smaller mean bias for the medians needs further study.

Although the results of our investigation and also that of Henry's (4) leads one to believe the mean rather than the median of watch times should be declared official time, the advantage is quite small. It is very doubtful whether the additional labor involved in computing the mean is justified under most circumstances.

There have been other instances of the application of questionable techniques in averaging scores in sports. For example, in a recent "Big Ten" championship gymnastics meet, the two extreme scores of the five judges

were discarded and the sum of the remaining three constituted a participant's official score. This was done to avoid individual prejudices, but the advisability of discarding these extreme scores is questionable. Probably a better plan would be to include all five in computing the official score.

Summary

Present practices in stop-watch timing require the median of three watch times to be declared "official" time. Theoretical evidence is available that the standard error of the median is larger than the standard error of the mean when sampling is from a single, normal population. The relationship between the standard error of the mean and standard error of the median may be expressed as follows:

$$\frac{\sigma Md}{\sigma M} = 1.160 \text{ for } N = 3$$

$$\frac{\sigma Md}{\sigma M} = 1.198 \text{ for } N = 5$$

An investigation was designed to determine if, under actual race conditions, the same relationship between the mean and median time exists. Two independent experiments yielded the following results, which are close to the theoretical values.

$$\frac{\sigma Md}{\sigma M} = 1.09 \text{ and } 1.17 \text{ for } N = 3$$

$$\frac{\sigma Md}{\sigma M} = 1.20 \text{ and } 1.16 \text{ for } N = 5$$

REFERENCES

1. Abraham, H., "The Answer Is in the Negative," *Amateur Athlete*, 22:8-9+, 1951.
2. Cureton, T. K. and D. E. Coe, "An Aanalysis of Errors in Stop Watch Timing," *Research Quarterly*, 4:94-109, 1933.
3. Fetter, C. H., "A New Way of Splitting Seconds," *Bell Telephone Quarterly*, 11:293-300, 1932.
4. Henry, F. M., "The Loss of Precision from Discarding Discrepent Data," *Research Quarterly*, 21:145-52, 1950.
5. Henshel, H. B., "Science Backs Harassed Officials," *Amateur Athlete*, 24:10+, 1953.
6. Hojo, T., "Distribution of the Median, Quartiles and Interquartile Distance in Samples from a Normal Population," *Biometrika*, 23:316-60, 1931.
7. Pearson, E. S. and N. K. Adyanthaya, "The Distribution of Frequency Constants in Small Samples from Symmetrical Population," *ibid.*, 20a:356-60, 1928.
8. Pearson, K., "Appendix to a Paper by Professor Tokishige Hojo," *ibid.*, 23:361-4, 1931.
9. Snedecor, G. W., *Statistical Methods*. Ames, Iowa: The Iowa State College Press, 1946, p. 485.
10. Warren, Ray, "Bulova Phototimer Sustains Officials," *Amateur Athlete*, 22:16-7, 1951.

(Submitted 8-5-53)

Acute Effects of Smoking on Physical Endurance and Resting Circulation[1]

PAUL A. PARKER
University of California
Berkeley, California

Abstract

The influence of smoking a cigarette or "pseudo-cigarette" on endurance times for a step test to exhaustion was compared for 19 male college students, using a balanced test-retest procedure. Blood pressures and pulse rates were markedly increased by smoking, but individuals classified as "tobacco-sensitive" by this test showed no greater change in endurance than "non-sensitives." The endurance of the total group was not significantly influenced by smoking. The pattern of cardiovascular changes caused by smoking were similar to the changes reported for epinephrine, suggesting that the total effect is vasodilation, which over-rides the cutaneous vasoconstriction.

IN RECENT YEARS there has been a revival of interest in the acute effects of tobacco smoking on physical performances and their related physiological processes. Disagreement of results has been the rule rather than the exception, particularly in large muscle activities. Karpovich and Hale, using a heavily loaded bicycle ergometer, reported a differential effect of smoking upon endurance, as eight of his subjects were not and five were affected by smoking (9). Henry and Fitzhenry, using a lighter ergometer workload, found no effect on oxygen intake or debt, or the rate of oxygen debt payoff (6). Juurup and Muido found no effect of smoking upon the oxygen intake with a heavier work load (8). Reeves and Morehouse, using a step-up exercise, found that smoking had no significant effect upon endurance, as well as no effect upon tests of speed, strength, or agility.

These negative findings are opposed by such studies as Cureton's, who found non-smokers to be superior to smokers in swimming (1). Also Steinhaus and Grunderman report a distinct superiority of non-smokers over smokers in cross country running (15). Dawson, however, points out that opposite results have occurred with marathon runners (2).

Many studies exist on the cardiovascular response to smoking; the consensus of these appears to be that smoking causes an immediate increase in systolic and diastolic blood pressures, and an increase in pulse rate, coupled with a peripheral vasoconstriction. Schilpp (13) has reviewed studies on these effects recently, particularly as they relate to heart rate and exercise.

---

[1]From the research laboratories of the Department of Physical Education, University of California, Berkeley. The writer is indebted to Dr. Franklin M. Henry for advice and help during the course of this investigation. The writer of this article wishes to state that the results of this research have no bearing whatsoever on the question of whether the continued use of cigarettes is harmful or non-harmful.

Since smoking effects may possible be influenced by the factors of prejudice or suggestion, it appears justifiable to view with some askance the studies that have neglected adequate control measures in this respect. This approach is even more necessary after the findings of Karpovich and Pestrocov, who showed the necessity of controlling these factors in studies concerned with the influence of ergogenic aids on exercises of an endurance nature (10).

Since some of the work loads used in previous studies may have been insufficient to reveal detrimental effects of smoking upon endurance, it seems desirable to conduct an investigation using a maximal large-muscle exercise. Another interesting possibility is to follow up the concept of Karpovich and Hale that some individuals are tobacco-sensitive, classifying them on the basis of their cardiovascular response to smoking. These individuals should be the ones to experience the greatest decrease in endurance as a result of smoking. A subsidiary problem concerns the interpretation of calculations in the resting circulation as revealed by pulse rate and blood pressure changes. Henry (4) has recently investigated this problem with respect to the theoretical changes produced by athletic conditioning; it should be of value to determine whether the impact of smoking on the resting circulation is consistent with his theory.

Method and Procedure

Nineteen male college students were used in the investigation. The age range of the subjects was from 19 to 25 years; the mean age was 22.2 years. Twelve were regular smokers, seven were non-smokers. All of the subjects were volunteers, either physical education majors or individuals who were actively participating in physical activity classes. Their willingness to work to exhaustion is shown by the fact that the average amount of work done by the subjects was 154,556 foot-pounds. Previous to this investigation, the greatest amount of work reported in smoking experimentation was 48,928 foot-pounds (9).

The experimental design involved a balanced test-retest procedure. The experiment was divided into three phases: a resting phase, a smoking or pseudo-smoking (control) phase, and an exercise phase. Half of the subjects smoked on their first test and half did the control experiment in their first test; this was determined by random assignment. External factors, such as pre-test activity and food ingestion, were controlled insofar as possible. Test-retest time, room temperature, instructions, etc., were also standardized. Blood pressure and pulse rate determinations were taken during the resting phase and continued at two-minute intervals during the smoking or control phases. During the rest period, the use of the exercise equipment was explained and demonstrated, and information was given to the subject concerning the procedure of the test. The smoking or pseudo-smoking phase was started when resting level cardiovascular measurements were consistent.

The smoking phase consisted of smoking one king-size cigarette down to three-eighths of an inch, or smoking from a hot air device that imitated a

cigarette. Both the smoking and control periods were accomplished on a time schedule with the subject inhaling once every 20 seconds upon command. The time schedule allowed some degree of standardization for the amount of smoke ingested during the smoking period, and helped considerably in heightening the deceptive aspects of the apparatus used to simulate a cigarette.

The metronome, used to maintain a steady exercise rate, was then started, and the subject was given a brief standardized "pep talk" before starting the exercise. As a last bit of motivation, the subject was told that a blood pressure check would be taken at the end of the exercise, and that this post-exercise measure would indicate whether or not he had given an extended endurance effort. The exercise was essentially the same as that used by Taylor in his test of exercise tolerance (16), that is, a two-count step-up exercise. The subjects stepped up on an eighteen-inch stool to a standing position at the rate of 46 steps per minute. They used a horizontal stall bar, placed 6 feet from the floor, to help in stepping up. One complete step consisted of using the right arm and right leg (or left arm and leg) to ascend the stool to a full vertical position, and then back down. The pack technique developed by Taylor was not used however, as in the experience of Henry[2] the addition of ten-pound weight increments does not noticeably improve the endpoint of the exercise, and could exert a variable psychological effect upon different subjects. Furthermore, since the subjects exercised in cadence with a metronome, it was readily apparent when each subject was becoming exhausted. The subject continued the exercise until exhausted, with no limit on the number of beats he missed. The missed beats were recorded and subtracted from the subject's total time.

Apparatus. The hot air device used to simulate a cigarette was developed after considerable preliminary experimentation. It consisted of a 660-watt electric heater element, mounted inside an asbestos-covered coffee can. Hot air from this unit was sucked in by the subject through a short length of pyrex tubing, then rubber tubing. A wooden mouthpiece, bored with a #57 drill hole, was attached to the rubber tubing. This combination approximated the suction resistance of a cigarette. It was necessary to keep the tubing short in length so that the air was still quite hot when it reached the subject's mouth. He could only see the outside of the box containing the unit, the rubber tubing, and the mouthpiece. For smoking cigarettes, an identical external tube arrangement and mouthpiece came out of the box; inside, however, it was connected to a cigarette holder.

As with other control devices in smoking experimentation (7), it was necessary to blindfold and place a nose clip on the subjects, since with the olfactory and optic senses excluded the problem of imitating cigarette smoking is greatly simplified. The nose clip and blindfold were also worn for cigarette smoking, and the subjects were misled into thinking they were

---

[2]F. M. Henry. Personal communication.

comparing two kinds of cigarettes. The device was successful, although its effectiveness appeared to depend to a considerable degree upon suggestions; the subjects were deceived by false impressions as well as the degree to which the device imitated a cigarette.

Results and Discussion

Endurance. A comparison of the effect of smoking upon endurance, by use of the t-ratio method of determining significance, was made between control and endurance times. The total mean difference between the smoking and control tests was 1.13 minutes, the endurance being less under the smoking conditions. The standard error of this difference was 1.05, which with 18 degrees of freedom, gave a t ratio of 1.07, a value that is far too small to be significant at the 5-per-cent level of confidence. The results make untenable the hypothesis that the immediate effects of smoking have a detrimental effect upon maximal performance. Since the control of major variables was greater than would be found in a real life endurance situation, the chances are almost nil that smoking a cigarette just before an athletic event has any important effect of "cutting the wind", i.e., of decreasing large-muscle endurance.

In order to determine to what extent a practice effect was present during the endurance retest, the first test endurance times were compared with the second test times. The t-ratio value was found to be 1.75, which is below the 5-per-cent level of confidence, indicating that no significant practice effect occurred during the retest. In fact, the subjects were quite consistent in their test-retest endurance times, as a Pearson product-moment method of correlation gave $r = .882$.

To ascertain whether or not smoking had a detrimental effect upon the tobacco-sensitive subjects (i.e., the subjects who showed a relatively greater heart response to smoking), the group was divided on the basis of their heart rate response to smoking. The endurance times of the smoking and control tests of the half of the subjects with the greater pulse rate increase from smoking were compared. The t-ratio value was found to be 0.76. The half of the group with the smaller pulse rate increase were also compared in a similar manner and the t value was found to be 0.80. Neither of these values was significant at the 5-per-cent level of confidence. A comparison between the amount of change in endurance in the first (greater pulse rate increase) group with the amount of change in the second (lesser pulse rate increase) group was found to be $t = 0.35$. Again this value is not significant, indicating that tobacco sensitivity is not significantly related to the effect of smoking on endurance. A further check of this finding was made by correlating the heart rate changes from smoking with the endurance times changes from smoking for all subjects. This value was found to be $r = 0.154$, which is not significantly different from zero. We therefore have no evidence in this study that tobacco sensitive individuals have a greater change in endurance due to smoking than non-sensitive individuals.

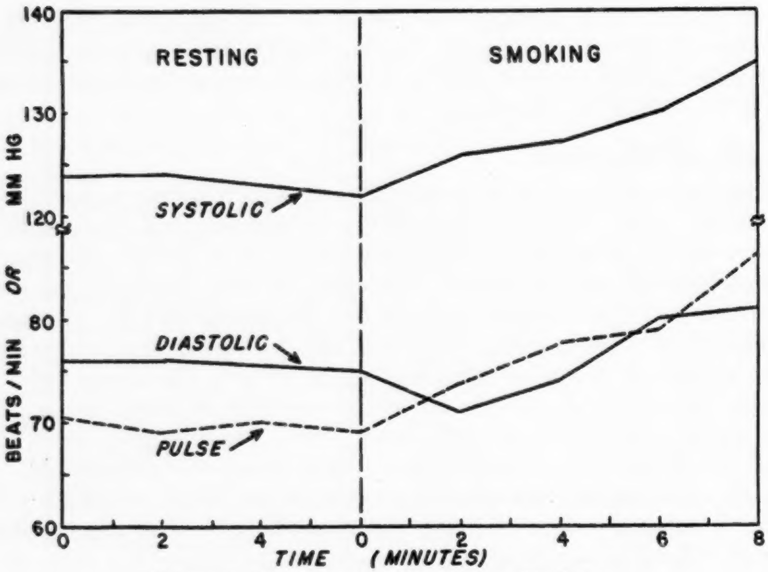


FIGURE I. Cardiovascular Response to Smoking

In view of the completely negative results to this point, there seemed to be no reason for further analyzing the data on effects of smoking upon endurance.

Cardiovascular Changes. To investigate the problem of the relative sensitivity of various cardiovascular measures to the effects of smoking, t ratios were first calculated for the effects of smoking on these measures. The results are shown in Table 1. As may be noted, all of these measures show significant increases at better than the 1-per-cent level of confidence; these results agree with other studies on the effects of smoking upon the cardiovascular system. A comparison between the resting and pseudo-smoking periods did not reveal any changes that were significant at the 5-per-cent level of confidence, indicating that the control device did not produce any important physiological effects and may be considered successful in this respect. These results are shown in Table 2. The nearly significant drop in diastolic pressure is in the direction expected, since the control device delivered hot air, and could have some vasodilatory influence.

Interpretation of Cardiovascular Changes. Previous investigations by Short and Johnson have demonstrated a remarkable parallelism between the

TABLE 1

Changes in Cardiovascular Measures due to Smoking

Measure	Means Pre-Smoking	Changes due to Smoking	Standard Error of the Diff.	t Ratio
Pulse Rate	69.67	17.49	3.05	5.73
Systolic B.P.	122.85	9.67	1.81	5.34
Diastolic B.P.	75.98	5.27	1.51	3.49
Mean B.P.	99.31	7.47	1.57	4.75

TABLE 2

Changes in Cardiovascular Measures due to "Smoking" Hot Air

Measure	Means Pre-Smoking	Changes due to Smoking	Standard Error of the Diff.	t Ratio
Pulse Rate	69.34	1.19	1.11	1.07
Systolic B.P.	123.30	—1.52	1.35	1.12
Diastolic B.P.	74.36	1.99	1.00	1.99
Mean B.P.	98.83	0.24	1.30	0.23

effect of epinephrine (adrenaline) injections and tobacco smoking (14). Measurements of skin temperatures, oral temperatures, blood pressure, pulse rate, and blood sugar during both smoking and epinephrine tests displayed such similarity that the investigators concluded that the characteristic physiological effects of smoking are due to the increased output of epinephrine. Ralston et al. have pointed out that although epinephrine causes cutaneous vasoconstriction, the total vascular effect is known to be vasodilation and increased cardiac output (11). They have shown that evidence of this dilation can be seen in the cardiovascular dynamics when the blood pressure and heart rate changes are interpreted by the E/W method of interpreting cardiovascular measurements. (E/W is the ratio of arterial elasticity to peripheral resistence) (11). Henry has recently expanded this method, and has shown the importance of two other cardiovascular calculations that can be obtained from the heart rate and blood pressure measurements (4). These are EV_s, defined as the product of the arterial volume-elasticity coefficient E and the stroke volume of the heart V_s, and WV_s, defined as the product of the vascular resistance coefficient W and the stroke volume.

The essence of the Henry study, as it applies to the present one, is that even though the stroke volume is not measured, the *pattern* of changes in the various blood pressures and EV_s, WV_s and E/W are uniquely different depending on what changes have occurred in stroke volume, heart rate, arterial elasticity, and vascular resistance. He has prepared a table to show several possible patterns and has demonstrated that the cardiovascular effects of athletic training fit in with one specific pattern. Similarly, the cardiovascular changes due to smoking also fit one specific pattern, as will be shown. Thus it is possible to substitute observed cardiovascular values (i.e., blood pressure and pulse rate) in the specific "smoking pattern" of cardiovascular response and estimate changes in other important circulatory functions that are not directly available (i.e., V_s, W, and E).

The pattern of changes due to smoking has been determined for observed and estimated values using the Henry method of cardiovascular analysis, and may be seen in Table 3. Changes in volume-elasticity (E) must be excluded as a possibility, since, being anatomical, they would require weeks or months. They could not occur within the 7-minute smoking period. If the value of E is assumed to be 1.20 (see Henry's Table 4) and the known pulse rate values are transposed into beats per second (R), and E is considered constant, then the changes in stroke volume and other important

measures may be derived algebraically. These values, computed from the formulae given by Henry (4), are shown in Table 4. Of particular interest are the values for V_s, which show an increase in stroke volume. It does not seem to be possible to account for the observed blood pressure and pulse rate changes without postulating an increase in stroke volume. Previous assumptions have been that stroke volume decreases upon smoking (3, 13).

TABLE 3

Changes in the Cardiovascular System due to Smoking

Measure	P_s	P_d	P_m	P	EV_s	WV_s	E/W
Pre-smoking	122.9	75.8	99.4	47.1	94.2	85.6	1.10
Seven minutes after smoking	132.5	81.1	106.8	51.4	102.8	73.5	1.40
Changes due to smoking	9.6	5.3	7.4	4.3	8.6	—12.1	0.30

TABLE 4

Calculated Physiological Changes due to Smoking

Measure	R (*beats/sec.*)	E	W	E/W	V_s (*cc/beat*)
Before smoking	1.161	1.200	1.090	1.101	78.5
After smoking	1.453	1.200	0.887	1.345	85.6
Changes due to smoking	0.292	0.000	—0.203	0.235	7.1

Also of significance is the fact that a considerable *decrease* in vascular resistance occurs as a result of smoking in spite of cutaneous vasoconstriction. This same pattern of changes is revealed as a result of epinephrine, by calculations from the figures given by Ralston et al (11) for the cardiovascular effects of this drug. The results therefore support the conclusions of Short and Johnson as to the "epinephrine effect" of smoking.

Since a variety of cardiovascular measures were available as indicators of the physiological effect of smoking, calculations were made in order to determine what particular measure would yield the most sensitive test of the effects of smoking. Making use of the *t* ratio in the manner developed by Henry and Berg (5), it was found that the observed, rather than estimated, cardiovascular measures gave the best test for sensitivity to smoking. The most sensitive indicators were the pulse rate ($t = 5.73$) and the systolic blood pressure ($t = 5.34$), as may be seen in Table 1.

Summary and Conclusions

Using a step test exercise to exhaustion, 19 male subjects of college age were tested for endurance time following a smoking period and again following a period of "smoking" a hot air device, using a balanced test-retest procedure. Records of the pre-smoking and smoking pulse rates and blood pressure were taken for the test and retest. From the results of the study it may be concluded:

1. The acute effects of smoking one cigarette show no statistically significant deleterious effect upon endurance as measured by a maximal step test.

2. Smoking one cigarette produces a significant increase in the pulse rate, systolic blood pressure, and diastolic blood pressure.

3. The most consistent index of the immediate physiological effects of smoking is the pulse rate increase.

4. Defining tobacco sensitivity as an above-average increase in heart rate due to smoking, the endurance of tobacco-sensitive individuals is no more influenced by smoking than is the endurance of less sensitive individuals.

5. The total vascular response to smoking is one of dilation, rather than constriction. This effect is similar to that of epinephrine injections, including a probable increase of stroke volume of the heart.

6. Subjects tested by use of a maximal endurance test of the step-up type display a high reliability of performance ($r = .882$).

REFERENCES

1. Cureton, T. K., Jr., Factors Governing Success in Competitive Swimming. *Swim. Pool Data and Refer. Ann.*, **4**: 45-49 (1936).
2. Dawson, P. M., *The Physiology of Physical Education*. Baltimore: Williams and Wilkins Co., p. 509 (1935).
3. Henderson, Y., H. W. Haggard and F. S. Dolley, The Efficiency of the Heart and the Significance of Rapid and Slow Pulse Rates. *Am. J. Physiol.*, **82**: 512-524 (1927).
4. Henry, F. M., Influence of Athletic Training on the Resting Cardiovascular System, *Research Quarterly*, **25**: 28-41 (1954).
5. ——— and W. E. Berg, Physiological and Performance Changes in Athletic Conditioning, *J. App. Physiol.*, **3**: 103-111 (1950).
6. ——— and J. R. Fitzhenry, Oxygen Metabolism of Moderate Exercise, with Some Observations on the Effects of Tobacco Smoking. *J. App. Physiol.*, **2**: 464-468 (1950).
7. Hull, C. L., The Influence of Tobacco Smoking on Mental and Motor Efficiency. *Psych. Mono.*, **33**: No. 3 (1924).
8. Juurup, A., and L. Muido. Acute Effects of Cigarette Smoking on Oxygen Consumption, Pulse Rate, and Blood Pressure in Working Organisms. *Acta. Physiol. Scandinav.*, **11**: 48-60 (1946).
9. Karpovich, P. V., and C. J. Hale. Tobacco Smoking and Physical Performance. *J. App. Psych.*, **3**: 616-621 (1951).
10. ———and K. Pestrocov. Effect of Gelatin Upon Muscular Work in Man. *Am. J. Physiol.*, **134**: 300 (1941).
11. Ralston, H. J., O. E. Guttentag and E. Ogden. Evaluation of the Ratio of Aortic Rigidity to Peripheral Resistance. *Proc. Soc. Exp. Biol. and Med.*, **53**: 154-156 (1943).
12. Reeves, W. E., and L. E. Morehouse, Acute Effect of Smoking Upon the Physical Performance of Habitual Smokers. *Res. Quart.*, **21**: 245-248 (1950).
13. Schilpp, R. W., A Mathematical Description of the Heart Rate Curve of Response to Exercise, with Some Observations on the Effects of Smoking. *Res. Quart.*, **22**: 439-445 (1951).
14. Short, J. J., and J. H. Johnson, Relationship of Tobacco Smoking to Epinephrine. *J. Lab. and Clin. Med.*, **24**: 590 (1939).
15. Steinhaus, A. H., and Florence Grunderman, *Tobacco and Health*, New York; Association Press (YMCA), (1948).
16. Taylor, C., A Maximal Test of Exercise Tolerance. *Res. Quart.*, **15**: 291-302 (1944).

(Submitted 8-1-53)

An Inexpensive Gravity Reaction Time Device

A. T. SLATER-HAMMEL

Indiana University
Bloomington, Indiana

Abstract

This article describes a simple and inexpensive hand reaction time device which should help meet the needs of institutions with limited research budgets. A detailed description of construction procedures and use of the device is presented.

ONE NOTICEABLE trend in experimental research in physical education has involved the use of increasingly complex apparatus. Although this complication of apparatus has served to facilitate and broaden our experimental procedures, it has also increased enormously the cost of experimentation. This factor of cost has acted not only as a deterrent to the establishment of experimental laboratories in departments of physical education, but it appears to have left many young professional workers with the impression that most experimental work is *necessarily* an expensive business. But the establishment of a laboratory and the conduct of experimental research need not be excessively expensive. With a little ingenuity and a few simple tools, it is possible to construct many inexpensive pieces of apparatus which are completely adequate for a variety of experimental problems.

It is the purpose of this paper to describe a simple and inexpensive gravity device for the measurement of hand reaction time. The construction requirements are so elementary that the device can be duplicated by anyone whose talents include the operation of a saw and drill.

Principle and Description of Apparatus

The operation of a gravity reaction time device is based upon the fact that a body falling under the force of gravity moves with constant acceleration. Using this fact, reaction time can be measured under two conditions: (1) having the subject stop a falling body, or (2) having the subject move away from the path of a falling body.

The first procedure has been used in at least two reaction time studies (1, 2). In each case, the device used was a form of Atwood's machine in which a counter weight served to reduce the velocity of a falling body. With this device, the subject was simply required to stop the falling weight. The subject's reaction time was determined by the distance the weight fell before being stopped.

The second procedure does not appear to have been commonly used to measure reaction time. This is rather surprising because the apparatus requirements are much simpler than in the first procedure. All that is needed is a method whereby the reacting body part can be moved closer to or farther

away from the starting point of a falling body. To measure hand reaction time, for example, it is only neoessary that a method be available for determining how close a subject can bring his hand to the starting point of a falling body without being struck by the falling body. Having found this distance, the hand reaction time is determined by the formula $t = \sqrt{d/\frac{1}{2}g}$, where t = reaction time, d = distance of hand in centimeters from the starting point of the falling body, and g = 980 centimeters.

A device constructed for the measurement of reaction time by this procedure is illustrated in Figure I. Briefly, the device consisted of a 30 x 23 x $\frac{3}{4}$-inch wooden base upon which a $\frac{3}{4}$ x $\frac{1}{8}$-inch angle iron framework was mounted. The framework consisted of two 24-inch vertical bars, a horizontal bar mounted across the top of the vertical bars, and two angle supports. A 12 x 6 x $\frac{1}{16}$-inch metal plate was mounted flush with the bottom of the horizontal bar. At the back and bottom of the plate, a $\frac{3}{8}$-inch length of $1\frac{1}{4}$-inch diameter tubing was located. Screws projecting to the back of the vertical bars served as supports for the hand rest, which consisted of a 29 x $\frac{1}{2}$ x $\frac{1}{2}$-inch metal bar. The arrangement was such that the hand rest could be raised or lowered at will.

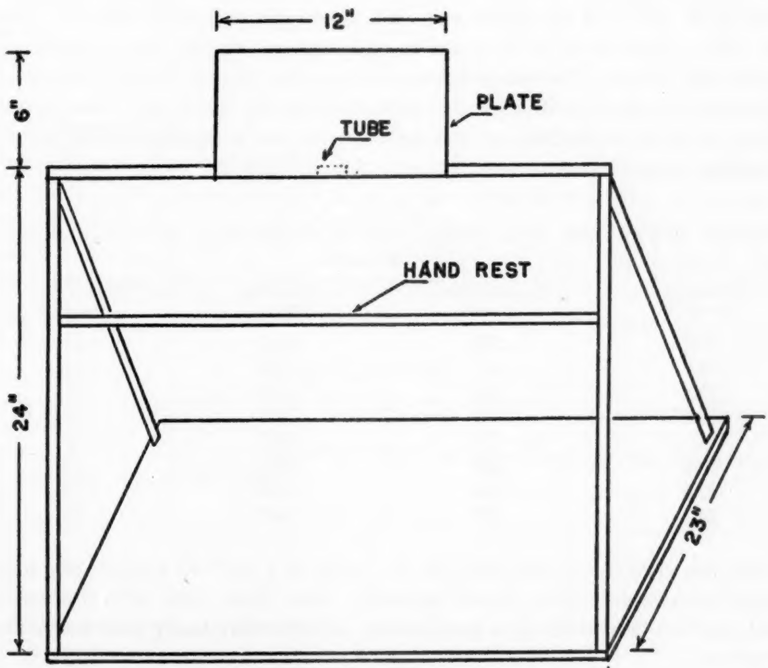


FIGURE I. Diagram of Gravity Reaction Time Device

To obtain a device which would measure reaction time over a range of .15 to .30 seconds and in .01 second intervals, it was necessary to locate the screw supports along the vertical bars so that the hand rest would be adjustable over specified distances from the bottom of the metal plate. These distances were determined by the general formula $d = \frac{1}{2}gt^2$, where $d =$ the distance a freely falling body drops, $g = 980$ centimeters, and $t =$ the time of drop. Table 1 summarizes the various distances at which the hand rest could be located from the bottom of the plate and the corresponding reaction time measures. The distances presented in Table I include an allowance of 1 centimeter for the fact that a subject must place his fingers on top of the hand rest.

Application of Apparatus

To measure hand reaction time, the subject is seated before the apparatus with the fingers of one hand resting on the hand rest and directly below the tube. The height of the chair should be adjusted so that the subject can sight directly across the bottom of the plate and tube. Upon the completion of these adjustments, the experimenter stands to one side of the apparatus and places his hand behind the plate, which serves as a shield. With his hand in position, the experimenter holds a small object—a marble, a small rubber ball, etc.—in the tube and just above the subject's line of vision. After the presentation of a suitable preparatory signal, the experimenter releases the object. The subject, upon seeing the falling object, attempts to avoid being struck by withdrawing his hand from the hand rest. The subject's reaction time is determined by the distance he can bring his resting hand to the bottom of the plate without being struck (Table 1).

TABLE 1
Distance of Hand Rest from Starting Point of Falling Body and Corresponding Time Measure

Distance (*in cm.*)	Reaction time (*in sec.*)	Distance (*in cm.*)	Reaction time (*in sec.*)
12.0	.15	26.9	.23
13.5	.16	29.2	.24
15.2	.17	31.6	.25
16.9	.18	34.1	.26
18.7	.19	36.7	.27
20.6	.20	39.4	.28
22.6	.21	42.2	.29
24.7	.22	45.1	.30

Since the gravity reaction time device involves a starting postion and a type of hand movement which varies markedly from those used with the conventional reaction time device,[1] a preliminary comparative study was undertaken.

[1]The conventional reaction time device includes a Standard Electric Clock which measures time in hundredths of a second, a light stimulus, and a telegraph key or similar device for operation by the subject.

In this investigation, the reaction times of 25 physical education majors at Indiana University were taken with each device.

In measuring reaction time with the gravity device, the hand rest was started from the lowest position and progressively moved upward. The subject was given five trials at each hand rest position. The level at which the subject's hand was struck at least four times determined his reaction time. Reaction time measures with the conventional device were based upon an average of 25 trials. In both procedures, the subject was permitted to practice until it became evident that he understood the task required of him.

Table 2 summarizes the reaction times obtained with each device. The slight difference between the two values was not statistically significant. The correlation between the two reaction time measures was found to be .842. In view of the fact that test-retest measures with the conventional reaction time device on these same subjects gave correlations ranging from .853 to .870, it would appear that the reaction times obtained with the gravity device are comparable to those obtained with the conventional device.

TABLE 2
Comparison of Reaction Time with Gravity and Conventional Reaction Time Device

Points of Comparison	Gravity Device	Conventional Device
Mean (*sec.*)	.200	.201
Standard Deviation (*sec.*)	.026	.018

Cost of Apparatus

The total cost of materials for the gravity reaction time device was approximately $6.00. It is to be emphasized, however, that the selection of materials described was simply a matter of using whatever was conveniently at hand. Actually, there is no reason why the framework and plate could not be constructed of wood. This would undoubtedly reduce the cost of material as quoted above.

REFERENCES

1. Burpee, R. H. and W. Stroll, Measuring Reaction Time of Athletes, *Research Quarterly*, **7**: 110-118 (1936).
2. Elbel, E. R., A study of response time before and after strenuous exercise, *Research Quarterly*, **11**: 86-95 (1940).

(Submitted 8-23-53)

A Study of Tests of Kinesthesis [1]

VERNON R. WIEBE

Tabor College
Hillsboro, Kansas

Abstract

Twenty-one tests of kinesthesis were administered to 15 college varsity men and 15 college men who had never lettered in high school or college varsity sports. Test reliabilities, validities, and interrelationships were first determined. "t" ratios were computed to determine the statistical significance of the obtained differences between varsity and non-varsity men. Fifteen of the tests have reliability coefficients which would recommend them as useful testing instruments. There was a kinesthetic difference in favor of the varsity men.

KINESTHESIS (*Gr. to move + perception*) is the position sense. Its receptors are located in the muscles, tendons, and joints. Through it the individual perceives bodily tensions and movements on the basis of what he is doing or in terms of their relation to some contemplated mode of behavior. Adequate stimulus for its receptors is the mechanical pressure and state of strain in the muscles, tendons, and joints resulting from bodily tensions brought about by bodily position.

The relationship of kinesthesis to physical education is pointed out in the following six statements. The first three statements are presented by the writer on logical and experimental grounds alone. The last three represent classical experimentation in the area of kinesthesis. Kinesthesis, then, is related to physical education because:

1. The "position sense," kinesthesis, is the sense of the material with which physical education principally concerns itself (i.e., the use of muscles in performing motor activities).

2. The functions ascribed to kinesthesis—co-ordination of body movements, development of skills, locomotion, posture, body control, manipulation, balance, and appreciation of weights and forces—are important elements in the teaching of physical skills.

3. The components ascribed to kinesthesis—perception of movement, tension or resistance, position, space perception, balance, relaxation, and effort—are familiar and oft-used concepts in physical education.

4. McCloy (12), in a preliminary study of the factors of motor educability, determined kinesthesis to be one of these factors.

5. Bass (2), in an analysis of the components of tests of semi-circular canal function and static and dynamic balance, found the second factor to be kinesthesis, as well as finding a loading of this factor in almost all of her proposed balance tests.

---

[1]This study was made in partial fulfillment of the requirements for Masters' degree at the State University of Iowa, 1951.

Balance has been generally considered as an element of motor skill or as a fundamental skill in the field of physical education. The emphasis which the Swedish balance beams received in former time, the place held by dancing at present, and recent studies in motor skills are evidence of this. (2, *p. 33*)

6. Phillips (15), in the introduction to his study of the relationship between certain phases of kinesthesis and performance during the early stages of acquiring two perceptuo-motor skills, says:

Coachs and psychologists seem uniformly to agree, principally on *a priori* grounds, that efficient kinesthesis is essential to refined motor performance. Although not as conclusively as might be expected, studies do indicate a tendency for the more skilled performer of a specific manual task to test higher on traditional kinesthetic tests than does the less skilled individual. (15, *p. 571*)

The Two Problems

Thirteen tests of kinesthesis were selected from nearly one hundred proposed tests suggested by 24 experimenters (1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23). These were selected on the basis of simplicity and promise. Eight others were constructed by the investigator, using ideas presented in the literature.

The first problem involved an investigation of the tests themselves, determining test reliabilities, validities, and interrelationships. The second problem involved an investigation of the relationships between the tests considered statistically adequate and athletic ability. Finally, a short predictive battery was perfected.

Review of Literature

Lepley (10) found reliability coefficients of .28 to .77, determined by the average intercorrelation among the separate trials permitted on each test. Fisher (5) reports reliability coefficients of .57 to .93 determined by correlating the odd numbered trials with the even numbered trials of each test, and correcting it by the Spearman-Brown Prophecy Formula. Young (23) obtained reliability coefficients of .73 to .93 using this same method. Phillips (15) used the test-retest method and found reliability coefficients of .59 to .84 on five tests, but found his other five tests to have reliabilities too low for further use.

Young (23) reports validity coefficients of .037 to .869 between each test and composite score of 18 tests.

Fisher (5), Phillips (15), and Young (23) report low intercorrelations among the tests they used.

Kerr and Weinland (7) found athletes superior to non-athletes in duplicating given pressures on the back and leg and the grip dynamometers. Nichols (14) reports essentially a zero correlation between ability to duplicate a given pressure on the grip dynamometer and the subject's best score for 18 holes of golf. Taylor (18) found a low but positive relationship between kinesthesis and success as a basketball player.

Young (23) found that three of her tests correlated .984 with the composite score of her 18 tests and could be used as a predictive battery.

Procedure

1. A battery of 21 tests of kinesthesis was formulated. Thirteen of these tests were taken from other studies, eight were constructed by the writer.
2. Thirty men students from the State University of Iowa were selected as subjects. Group A consisted of 15 varsity lettermen[2] selected by their varsity coaches as the best all-around athletes on their teams. Group B consisted of 15 men from the required physical education classes who had not earned a varsity letter in high school or college.
3. A preliminary study to gain proficiency in the techniques of administration was conducted. After this experience, the tests were revised into the final battery as it appears in this paper.
4. Each subject was tested individually by the writer.
5. The tests were first checked for reliability. This was done by correlating the scores of the two odd-numbered trials against the two even-numbered trials. (The total of the first three trials was later used as the final score.) These coefficients were corrected with the Spearman-Brown Prophecy Formula to secure an estimate of the reliability of the whole test. The tests which had a reliability coefficient of .70 were used in further computations.
6. Validity of the tests are computed by correlating the tests with the composite T-scores of the tests of kinesthesis which met the criterion of reliability. While the composite score criterion is certainly not infallible, it may be assumed that it offers one measure of kinesthesis since it is based on the scores of individual tests which were set up to measure kinesthesis in the terms of the definition of kinesthesis used in this study.
7. Intercorrelations were computed to determine the degree of relationship between the tests.
8. Kinesthetic difference between athletes and non-athletes were determind by computing the t statistic and from this finding the significance of the difference between the group means.
9. A short predictive battery was perfected by determining the best combination of four or less tests using the multiple correlation technique. A multiple regression formula was computed from the best combination of tests.

Analysis of Data

Table 1 shows the reliability coefficients and Spearman-Brown Prophecy Formula corrections for each of the tests. Fifteen of the tests meet our criterion of reliability.

Table 2 shows the validity coefficients for each of the retained tests. None of these tests exhibits a high coefficient of correlation with the criterion.

Table 3 shows the intercorrelations among the retained tests. All of the intercorrelations except three are very low. These three high intercorrelations could be due to the common movements found in each of these tests.

---

[2]Gymnastics, 3; Swimming, 3; Track, 2; Baseball, 2; Basketball, 2; Football, 2; Wrestling, 1.

Table 4 shows the group means and levels of confidence at which the difference between the group means is significant. With two exceptions, group A is superior to group B. In those exceptions, the difference in favor of group B is very small. With two exceptions (tests 19 and 21), the differences in favor of group A are significant at the ten per cent level of confidence or lower.

TABLE 1
Reliabilities

Test Number and Name	Alternate Halves r	Spearman-Brown r
1 Horizontal Space	.47	.57
2 Dynamometer—Duplicate	.16	.22
3 Vertical Space	.74	.81
4 Dynamometer—Relaxation	.83	.88
5 Stick Size	.55	.65
6 Dynamometer—Half Effort	.56	.66
7 Index Fingers	.47	.57
8 Arm Horizontal	.83	.88
9 Leg Raise	.84	.88
10 Separate Feet	.85	.90
11 Balance—Lengthwise	.89	.93
12 Balance—Crosswise	.89	.93
13 Free Throw	.65	.74
14 Basketball Pass	.87	.91
15 Baseball Throw	.90	.93
16 Baseball Fielding	.67	.75
17 Pedestrial Vertical Space	.76	.83
18 Knee Bend	.40	.50
19 Sargent Jump	.77	.83
20 Gross Movement	.85	.89
21 Walk Pathway	.63	.72

TABLE 2
Validities

Test Number and Name	r
3 Vertical Space	—.30
4 Dynamometer—Relaxation	—.06
8 Arm Horizontal	—.22
9 Leg Raise	—.43
10 Separate Feet	—.42
11 Balance—Lengthwise	.36
12 Balance—Crosswise	.21
13 Free Throw	—.34
14 Basketball Pass	—.47
15 Baseball Throw	—.48
16 Baseball Fielding	—.54
17 Pedestrial Vertical Space	—.49
19 Sargent Jump	—.14
20 Gross Movement	—.36
21 Walk Pathway	—.18

TABLE 3
Intercorrelations

Test Number and Name	3	4	8	9	10	11	12	13	14	15	16	17	19	20	21
3 Vertical Space															
4 Dynamometer—Relaxation	-.19														
8 Arm Horizontal	.00	.07													
9 Leg Raise	-.13	.29	.13												
10 Separate Feet	.19	.28	-.10	.24											
11 Balance—Lengthwise	.01	.20	.74	-12.	-.01										
12 Balance—Crosswise	.34	.28	.23	-.11	.17	.60									
13 Free Throw	.11	.20	.41	-.08	.14	.29	.13								
14 Basketball Pass	.25	.41	-.00	-.01	.36	-.29	-.06	.26							
15 Baseball Throw	.11	.44	-.10	.05	-.15	-.17	-.06	.43	.24						
16 Baseball Fielding	.10	-.31	-.17	.41	.04	-.40	-.20	-.10	.15	.16					
17 Pedestrial Vertical Space	-.03	.36	-.03	.49	.25	-.25	-.21	.13	.10	.32	.25				
19 Sargent Jump	.73	-.48	.09	.05	.16	.13	.15	.28	-.12	-.19	.25	-.18			
20 Gross Movement	.35	-.13	.10	.23	.26	.01	-.16	-.15	-.08	.02	-.17	.13	.13		
21 Walk Pathway	.23	.17	-.18	-.20	.19	.18	.36	.19	.01	.21	-.24	.09	.08	.22	

TABLE 4
Differences Between Groups

Test Number and Name	Group	Mean	Level of Confidence (*per cent*)
3 Vertical Space	A	2.82"	100*
	B	2.78"	
4 Dynamometer—Relaxation	A	28.34#	.1
	B	37.14#	
8 Arm Horizontal	A	29.80°	80*
	B	28.67°	
9 Leg Raise	A	15.73°	10
	B	19.93°	
10 Separate Feet	A	9.36"	10
	B	10.64"	
11 Balance—Lengthwise	A	15.54s	1
	B	11.07s	
12 Balance—Crosswise	A	8.33s	10
	B	6.87s	
13 Free Throw	A	51.20"	.1
	B	64.60"	
14 Basketball Pass	A	11.93°	10
	B	15.47°	
15 Baseball Throw	A	10.53°	.1
	B	27.87°	
16 Baseball Fielding	A	9.47	10
	B	12.27	
17 Pedestrial Vertical Space	A	5.13"	10
	B	6.37"	
19 Sargent Jump	A	5.00"	80
	B	4.82"	
20 Gross Movement	A	38.87"	10
	B	45.60"	
21 Walk Pathway	A	10.33	30
	B	12.20	

*In favor of Group B.

Table 5 shows the multiple correlations of logical combinations as well as the multiple regression formula for the best combination.

TABLE 5

Multiple Correlations and a Simplified Formula for Predicting Kinesthesis

Variables	R	
$R_{0.3,9,11,20}$	.69	0 = composite "T" score of 15 tests
$R_{0.9,10,11,17}$	.69	
$R_{0.3,11,17,20}$	.66	
$R_{0.3,10,11,20}$	.64	Other numbers are actual test numbers.
$R_{0.3,9,11}$	.63	
$R_{0.9,10,17}$	.61	
$R_{0.3,10,11}$	.60	
$R_{0.9,17}$	.55	

Kinesthesis Score = .5 (Balance Lengthwise) — (Leg Raise) — .5 (Vertical Space) — .5 (Separate Feet) + 50

Conclusions

1. Fifteen of the tests of kinesthesis have reliability coefficients which would recommend them as useful testing instruments.
2. None of these exhibits a validity coefficient high enough to warrant its use as a single test.
3. The low intercorrelations between the tests indicates that there is no general kinesthetic sensitivity, but that there are probably numerous specific factors.
4. There is a kinesthetic difference in favor of athletes.
5. The best combination of tests to measure kinesthesis in college men is the Balance Lengthwise, Leg Raise, Vertical Space, and Separate Feet.

REFERENCES

1. BALL, ROBERT V., Objective Measure of Emotional Instability; Study of Kinesthetic Motor Learning Among A Group of Juvenile Delinquents, *Journal of Applied Psychology*, **9**: 226-245 (June 1929).
2. BASS, RUTH I., An Analysis of the Components of Tests of Semicircular Canal Functions and of Static and Dynamic Balance, *Research Quarterly*, **10**: 33-52 (May 1939).
3. BROWN, JUDSON S., E. B. KNAUFT, AND G. ROSENBAUM, The Accuracy of Positioning Reactions as a Function of Direction and Extent, *The American Journal of Psychology*, **61**: 167, 182 (1948).
4. FILER, PAUL WEBB, *The Relation Between Kinesthetic Judgment and Success as an Instrumental Musician.* Unpublished Master's thesis, The Pennsylvania State College, 1934.
5. FISHER, ROSEMARY B., *A Study of Kinesthesia in Selected Motor Movements.* Unpublished Master's thesis, State University of Iowa, 1945.
6. HOLWAY, ALFRED H., AND LEO M. HURVICH, On the Discrimination of Minimal Differences in Weight; I. A Theory of Differential Sensitivity, *The Journal of Psychology*, **4**: 309-332 (1937).
7. KERR, W. H., AND J. D. WEINLAND, Muscular Perceptivity as a Trade Test, *The Journal of Applied Psychology*, **17**: 550-558 (1933).
8. LAIDLAW, ROBERT W., AND MARY ALICE HAMILTON, The Qualitative Measurement of Apperception of Passive Movement, *Bulletin of Neurological Institute*, **6**: 145-156 (Jan. 1937).

9. LANGFELD, H. H., AND F. H. ALLPORT. *An Elementary Laboratory Course in Psychology.* New York: Houghton Mifflin Co. 1926.
10. LEPLEY, WILLIAM M., *The Effectiveness of Certain Procedures Designed to Discriminate Between Promising and Unpromising Candidates for Flight Training.* Unpublished paper, The Pennsylvania State College, 1940.
11. LEUBA, JAMES H., An Apparatus for the Study of Kinesthetic Space Perception, *American Journal of Psychology,* **20**: 370-373 (1909).
12. MCCLOY, CHARLES HAROLD, A Preliminary Study of Factors in Motor Educability, *Research Quarterly,* **11**: 28-40 (May 1940).
13. NATIONAL ASSOCIATION OF PHYSICAL EDUCATION FOR COLLEGE WOMEN, *Victory Through Fitness,* Workshop Report, University of Wisconsin, Madison, Wisconsin, 46-47 (June 24-30, 1943).
14. NICHOLS, ANNE, *The Psychomotor Factors That Are Associated With Ability To Play Superior Golf.* Unpublished Master's thesis, State University of Iowa, 1935.
15. PHILLIPS, BERNATH E., The Relationship Between Certain Phases of Kinesthesis and Performances During the Early Stages of Acquiring Two Perceptuo-Motor Skills, *Research Quarterly,* **12**: 3, 571-586 (Oct. 1941).
16. ROCHEMONT, MESNIL DE, Über Eine Dritte Komponete für die Wahrnehmung von Gliederbewegungen, *Zeitschrift für Biologie,* **84**: 522-534 (1926).
17. ROFFEL, GERTRUDE, Visual and Kinesthetic Judgments of Length, *American Journal of Psychology,* **48**: 331-334 (April 1936).
18. TAYLOR, WILLIAM JAMES, *The Relationship Between Kinesthetic Judgment and Success in Basketball.* Unpublished Master's thesis, The Pennsylvania State College, 1933.
19. THORNDIKE, E. L., A Note on Accuracy of Discrimination of Weights and Lengths, *Psychological Review,* **161**: 340-346 (1909).
20. VAN RIPER, C., An Experimental Study of the Japanese Illusion, *American Journal of Psychology,* **47**: 252-263 (April 1935).
21. WALKER, J. F., Pantograph as an Instrument for Studying Kinesthesis, *American Journal of Psychology,* **47**: 691-692 (Oct. 1935).
22. WEBER, C. O., The Properties of Space and Time in Kinesthetic Field of Force, *American Journal of Psychology,* **38**: 597-606 (Oct. 1927).
23. YOUNG, OLIVE G., A Study of Kinesthesis in Relation to Selected Movements, *Research Quarterly,* **16**: 277-287 (Dec. 1945).

APPENDIX

Description of Tests

All the tests are performed without the use of the eyes. Deviation from the preferred score is recorded. The score is the total of three trials.

Test 1: Horizontal Linear Space (9). A yardstick is placed horizontally before the seated S. S is instructed to look at the 18-inch mark and sense its position. S is then blindfolded and instructed to point to the preferred position. No practice trials are allowed. Score is recorded to the nearest one-fourth inch.

Test 2: Grip Dynomometer, Duplicate One-Half Total Strength. S is seated comfortably at a table. S grips a dynamometer with his dominant hand, steadying it with the other. S is told to grip his comfortable maximum. One-half of S's maximum is computed and used as the preferred score. S is then blindfolded and guided to sense the preferred score. S then attempts to duplicate the preferred score. No practice trials are allowed. Score is recorded to the nearest one-half pound.

Test 3: Vertical Linear Space (9). A yardstick is placed vertically before the seated S. S is instructed to look at the 16-inch mark and sense its position. S is then blindfolded and instructed to point to the preferred position. No practice trials are allowed. Score is recorded to the nearest one-fourth inch.

Test 4: Grip Dynamometer, Relaxation. S is instructed to apply tension "a" sensing this. From tension "a" he is instructed to grip tighter to tension "b." From tension "b" he is instructed to relax to tension "a." Tensions "a" and "b" are arbitrary, being left to the S's discretion. Score is recorded to the nearest one-half pound.

Test 5: Arm Kinesthetic Judgment of Size (9). S is seated before a table with the index fingers resting on the edge of a table. S is told to sense the length of a 12" stick. S is then instructed to separate his hands, stick is removed and S attempts to duplicate the length of the stick between his fingers. Score is recorded to the nearest one-fourth inch.

Test 6: Grip Dynamometer, Half Effort. S is seated comfortably at a table. S grips a dynamometer with his dominant hand, steadying it with the other. S is instructed to exert any comfortable pressure and sense it. S is then instructed to exert one-half the sensed pressure. A practice trial without blindfold is allowed to familiarize the S with test. Score is recorded to the nearest one-half pound.

Test 7: Arm Kinesthetic Judgment in Precision (18). S is seated at table. A T-shaped board is placed so that its long side is perpendicular to S's frontal plane. A piece of paper is tacked on each side of the board which is perpendicular to S's frontal plane. S is blindfolded and instructed to raise his elbows and bring his inked index fingertips together. Score is measured to the nearest one-sixteenth inch.

Test 8: Arms Side 90° (23). S is asked to stand and instructed to raise the arm sideward to the horizontal with the palm facing downward. Score is recorded in degrees.

Test 9: On Side, Leg Raised 20°. S is asked to die on his non-dominant side. S is shown a stick figure drawing with its leg raised at a 20° angle. S is instructed to duplicate the angle seen. Score is recorded in degrees.

Test 10: Pedestrial Kinesthesia of Size (18) . S is asked to stand erect with the heels touching. S is then instructed to separate his heels so that the inside of the heels are 12 inches apart. Score is recorded to the nearest one-fourth inch.

Test 11: Balance, Lengthwise (23) S is instructed to place his dominant foot lengthwise on a balance stick, raise his left foot from the floor and see how long he can maintain his balance without touching his free foot or any part of his body to the floor. S is given one preliminary trial and then blindfolded for the test. Score is recorded to the nearest one-half second.

Test 12: Balance, Crosswise (23). S is instructed to place the ball of his dominant foot crosswise on a balance stick, raise the other foot from the floor and see how long the balance can be maintained without touching the free foot or any other part of the body to the floor. S is given one preliminary trial and then blindfolded for the test. Score is recorded to the nearest one-half second.

Test 13: Basketball Free Throws, Duplicate. S is blindfolded and instructed to throw four consecutive free throws in the conventional two-handed underhand method. S is instructed to duplicate the first free throw (i.e., make all throws alike). Score is recorded in inches.

Test 14: Basketball Pass, 45° to Dominant Side. S's toes are placed against a toeboard (base of quadrant) with his medial line at the vertex of the quadrant. S is instructed to attempt to pass the basketball (two-handed chest pass) 45° toward his dominant side. Score is recorded in degrees read from the quadrant.

Test 15: Baseball Throw, 45° to Non-dominant Side. S's toes are placed against a toeboard (base of quadrant) with his dominant shoulder joint over the vertex of the quadrant. S is instructed to attempt to throw the baseball (overhand) 45° toward the non-dominant side. Score is recorded in degrees read from the quadrant.

Test 16: Baseball Fielding. The center of a target with eight concentric circles is placed on the dominant side at the belt line of the S. Pointer is fastened to S's dominant hand by means of rubber bands. S is blindfolded and S is positioned so that the pointer touches the distal edge of the largest circle when his arm is extended. S is guided to sense the position of his arm when the pointer is directly in the center circle. S is then

instructed to assume a shortstop's fielding position (knees slightly flexed with hands on knees). From the fielding position, S is instructed to become erect and touch the sensed position with the pointer. The smallest circle counts one point with a successive increase of one point for each ring. Missing the target counts nine points.

Test 17: Pedestrial Kinesthesia of Vertical Linear Space (18). S is shown a line drawn 14 inches above the floor and instructed to estimate the height of the line. S is then blindfolded and instructed to place the bottom of his sole on the top of the line (i.e., step onto the "step"). Score is recorded to the nearest one-fourth inch.

Test 18: Knee Bend (10). S is instructed to stand with his back to the Sargent Jump chart. Height minus ten inches is calculated as the preferred score. S is guided to sense the preferred score. S stands erect and then bends the knees attempting to duplicate the preferred score. Score is recorded to the nearest one-fourth inch.

Test 19: Sargent Jump, Duplicate. S is instructed to jump four consecutive comfortable heights. S is instructed to duplicate the first height (i.e., make all four jumps alike). Score is recorded to the nearest one-fourth inch.

Test 20: Gross Kinesthetic Movement (18). S is told to toe a starting line. S is instructed to sense the distance of an arc drawn five yards from starting line. S is blindfolded and instructed to toe the arc. Score is recorded in inches.

Test 21: Walking on a Path (23). S is told to sense the direction of a path (two parallel lines ten inches apart and 18 feet long marked on floor). S is blindfolded and told to walk ten steps between the lines without touching them. Score is the number of steps on the line or off the path.

(Submitted 6-12-53)

Rate of Learning in Relation to Spacing of Practice Periods in Archery and Badminton[1]

OLIVE G. YOUNG
University of Illinois
Urbana, Illinois

Abstract

The purpose of this study was to compare the effectiveness of learning between two distributions of practice in archery and badminton classes. Each activity was divided into two groups based on the distribution of practice used. The "t" test of the significance of the difference between means was used to make comparisons between groups relative to the effectiveness of the two practice distributions of learning as measured by gain in selected criteria. The results seemed to indicate that more effective learning in archery took place under relatively massed practice while in badminton wider distribution of practice produced more rapid learning.

THE QUESTION OF the length and frequency of practice sessions or class instruction periods has long been of interest to teachers of physical education as it has to teachers in many other fields. Although the majority of experimental evidence has shown that, in general, learning is more efficient when practice periods are distributed rather than massed there has been very little research to determine what specific patterns of distribution are most effective in learning motor skills.

The results of research on distributive and massed learning by such psychologists as Duncan (6), and Hilgard and co-workers (3, 4, 5, 9) have indicated that, in most learning situations, spaced practices result in more efficient or more rapid learning than massed practices. A recent study by Harmon and Miller (8) indicated that after initial learning has taken place greater spacing between practice periods produces more favorable effects on the learning process than continued massing of practices. Distributive practices have generally been shown to be more effective than massed practices irrespective of whether the number of practices under massed and distributive situations were equal or not. Duncan (6) found that the amount of learning resulting from distributive practice was as great as that resulting from massed practice, even though the former had less actual practice time.

There were no studies found which used typical class learning situations as the basis for investigation. Would results from studies of the distribution of practice in learning physical education activities under class conditions follow the same pattern as the results obtained from learning in selected psychomotor skills under the more widely used and more readily controlled

[1]This study was made in partial fulfillment of the requirements of the degree of Doctor of Philosophy in the Department of Physical Education in the Graduate College of the State University of Iowa, 1953.

laboratory conditions? With this question in mind the present study was undertaken.

Statement of Problem

The purpose of this study was to compare the rate of learning in four-day-a-week and two-day-a-week distributions of practice in college archery and badminton classes; this rate of learning was to be determined by the amount of gain in selected criteria. The effect of motor ability and sex on the rate of learning in the two practice distributions and in the two activities was also studied.

Procedure

The regular physical education classes in archery and badminton of the Division of Health and Physical Education of Mankato State Teachers College during the school year of 1951-52 were selected for this study. The archery classes were all taught by the writer and the badminton classes were all taught by one other instructor.

All of the archery classes were taught on an indoor archery range. It was necessary to restrict the shooting distance; the maximum distance possible was 20 yards. Because of the short distance, 24-inch indoor target faces were used rather than the regulation 48-inch targets. Registration in each class was limited to 12 students, and six were assigned to each target.

The badminton classes were all taught on the same three badminton courts in one gymnasium. The courts which were placed across the width of the gymnasium were two feet shorter than regulation, but were of regulation width. In the high clear test, used as one of the criteria for measuring improvement, the distance from the net to the rope was reduced from 14 feet to 13 feet six inches.

Because only three courts were available, the registration in the badminton classes was limited to 12. Plastic shuttlecocks were used in all of the classes participating in the study throughout the year.

The classes used in the study were regularly scheduled physical education activity classes which fulfilled the college requirement for graduation and were open to all students, both men and women. All classes met for a 40-minute activity period and ordinarily were scheduled to meet two days a week for a 12-week quarter. For this study it was possible to make arrangements for some of the classes to meet four days a week for half of the length of the regular quarter.

The classes in each activity were divided into two groups on the basis of the number of days a week which the class met. Group I, the two-day-a-week group, was made up of those classes meeting twice a week, and Group II, the four-day-a-week group, was composed of the classes meeting four times a week. It was not possible to arrange the schedule in such a way that all classes in Group I could meet on any two specific days, or that the Group II classes could all have an identical pattern. Consequently, there is some slight variation in the pattern of practice distribution between classes within each

group. The number of 40-minute class periods was the same for all groups.

Group I in archery was made up of four classes with a total of 35 students and Group II consisted of three classes and 28 students. Groups I and II in badminton were each composed of four classes with a total of 35 and 41 students respectively.

Lesson plans were made on the basis of 20 lessons, the maximum number possible, for both archery and badminton classes, but due to irregularities in the length of quarters and to vacations all classes were not able to complete all 20 lessons. The statistical analysis of the data gathered from the archery classes was based on 19 lessons, while the analysis of the data from the badminton classes was based on 16 lessons.

As the purpose of this study was to compare the rate of gain made by the two-day-a-week group with that of the four-day-a-week group in both archery and badminton classes, the lesson plans for each activity included provisions for testing throughout the series of lessons. On the basis of the scores derived from these tests, learning curves were drawn and tests of the significance of the difference in improvement between Group I and Group II were made.

In order to restrict the personnel of the two groups to beginners in both activities, each student was asked to indicate how much badminton or archery instruction he or she had had, or how often each had participated in the activity. Those who indicated class or individual instruction, or who had participated in the activity more than six times, were eliminated from the study but not from the class. The students in each class were also asked not to practice that particular skill outside of class during the quarter in which the class met. Their special co-operation was asked with respect to absences and, although some absences did occur, those cases were not discarded unless they exceeded four in number or included more than three testing periods.

It seemed advisable to use the same battery of motor ability tests to equate both men and women relative to their potential ability to learn physical education or motor skills. The Scott Motor Ability battery (14) was selected and administered to 265 women in the physical education activity classes. A modified Scott Motor Ability battery was administered to 185 men registered in regular physical education activity classes.

In order to use the Scott Motor Ability battery for the men as well as the women, it was necessary to find some means of measuring throwing ability for men other than the basketball throw, as the men were able to throw a basketball too far for indoor administration of the test. A shot-put with a 16-lb. shot was tried first. It was found that this test was too difficult to administer, even with indoor facilities, and that it did not parallel closely enough the basketball throw. A throw using a five-pound medicine ball was tried and the results were found to be highly comparable to the basketball throw for women. A reliability coefficient of .80 was obtained for the medicine ball throw when the best two of three trials were correlated. As exten-

sive work had been done on the validity of this battery for women it seemed reasonable to assume that similar results would be found for the men. The medicine ball throw was therefore used instead of the basketball throw when the battery was given to the men.

The motor ability battery was given to as many students as possible in order that T-scales might be constructed for this particular college. Each of the tests of the battery was T-scored separately for men and women, and the sum of the T-scores for the three tests was used to equate the archery the badminton classes. There was considerable overlap of the sum of the T-scores for men and women, with both men and women at the upper and lower limits of the total range.

In archery a daily record was kept of the number of ends of six arrows which were shot, the number of arrows hitting the target (hits), and the score for each end (score). All classes followed exactly the same plan with respect to the number of ends shot each lesson. The rate of learning in archery was based on the shooting average, which is herein defined as the mean score per end per person. A second learning curve was drawn using the percent of hits in one day's shooting as the basis of computing the learning rate (see Figure I).

In badminton three tests were selected from those described by Scott and French (14). These included the 30-sec. wall volley, the short serve, and the high clear. Because of the ease of administration and the minimum amount of time involved, the wall volley test was given seven times, the short serve test was given five times, and the high clear test, which is very time-consuming in administration, was given only three times during the series of lessons.

In both archery and badminton, the mean gain for each group was computed, and the difference in gain between the two groups determined. Comparisons were also made with respect to differences in gain made by the women and by the men in both groups, differences between men and women within groups, and between high and low motor ability both within and between groups.

Analysis of Data

The *T*-scores on the three motor ability tests were added and this sum was used to represent the motor ability score for each student. Table 1 indicates that there was no real difference in mean motor ability between the two groups in each activity. It also shows that there was no significant difference between the standard deviations of the two groups in each activity with respect to motor ability. The scores of all subjects in archery and badminton were combined into one distribution. For the purpose of this study, high motor ability was interpreted as all scores above 170 and low motor ability as all scores below 140. These two reference points were selected for classification purposes because they indicated approximately the 75th and 25th percentiles respectively.

TABLE 1
Range, Mean, and Standard Deviation of Motor Ability T-scores in Archery and Badminton

Class	Number	Range	Mean	SD
Archery				
Group 1	35	111.4	155.21	21.98
Group II	28	94.7	155.51	21.52
Badminton				
Group 1	35	121.7	157.59	22.15
Group II	41	94.5	157.83	21.06

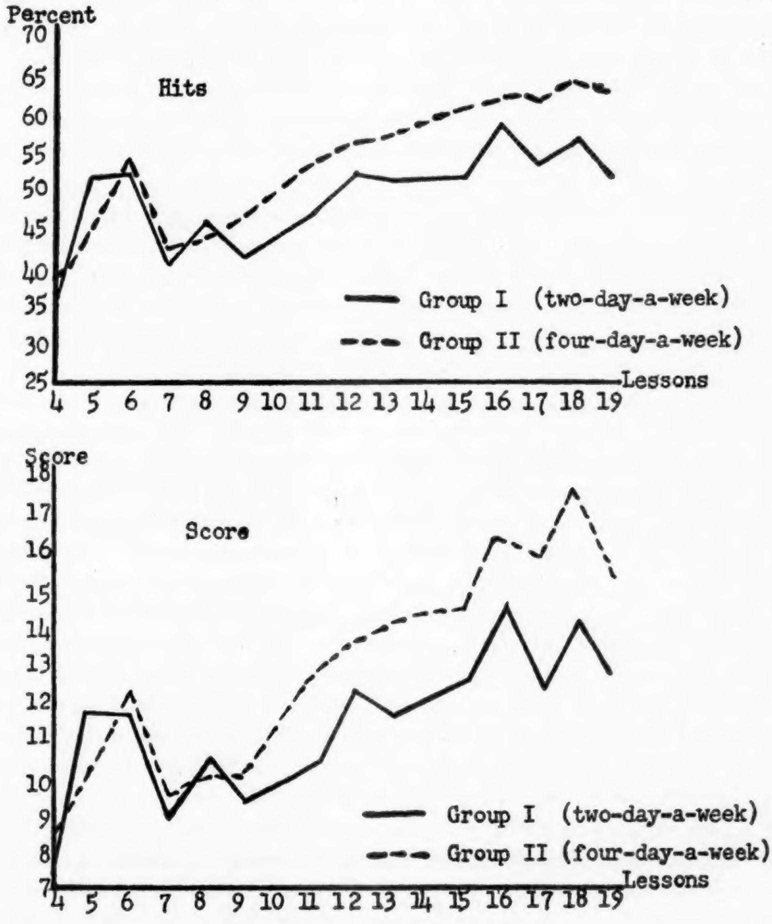


FIGURE I. Archery Learning Curves

ARCHERY

After each practice session, the mean score of the percentage of hits was determined for each individual in each class. As previously stated, the score

or shooting score refers to the total score for any one day divided by the number of ends of six arrows shot. The hits or percentage of hits was defined as the total number of hits divided by the total number of arrows shot.

From these daily averages, learning curves indicating progress were plotted for the two-day-a-week group and for the four-day-a-week group for both scores and hits. Both the scores and hits showed a sharp rise from lesson four, on which day scores and hits were recorded for the first time, through lesson six. On lesson seven the shooting distance was changed from 15 to 20 yards, which probably accounted for the marked drop in both scores and hits between lessons six and seven. From lesson seven on, the distance was held constant at 20 yards, except on lessons 16 and 18, when an equal number of ends was shot from both the 20-yard and the 15-yard distances. In lessons ten and 14, arrows were shot from both distances, but because these were special novelty shoots, no scores were kept. The learning curves for both scores and hits made sharp rises on the days when results were recorded for both 15 and 20 yards.

This rise was probably due to the greater accuracy with which the students shot at 15 yards. The marked drop on lesson 19 might have been due to the special nature of that lesson. Each student, alternating with a partner, shot just one arrow at a time, rather than all six consecutively. The object was to break balloons hung over the target. Scores were recorded, however, so that a mean score and percentage of hits could be determined.

Although both groups began at nearly the same level, Group I showed a slightly more rapid gain between lessons four and five and between lessons seven and eight, but from lesson eight on Group II showed a steadier and more rapid improvement than did Group I. In order to determine the amount of gain made in learning during the term of instruction, it was necessary to have an initial and final evaluation of performance. The inital value for the scores was determined by obtaining the mean score per end for lessons four, five, six, and seven combined. The final evaluation was obtained by finding the mean score per end for the combination of the last three lessons. This same procedure was followed to obtain initial and final percentages for the hits. This procedure considered initial scores on both the 15-yard and the 20-yard distances, as well as final scores on both distances. Such an average represents, in the opinion of the writer, a fairly accurate evaluation of initial and final ability in archery.

In order to be certain that the gain between the initial and final achievement was real and not just due to chance, the t test[2] of the significance of the difference between means was made for both groups for scores and hits. With only one exception, the difference between initial and final scores as well as initial and final percent of hits, for the groups as a whole and for men and women separately, was significant at the 1-per-cent level of confi-

[2] $t = \frac{M_{diff}}{\sigma M_{diff}}$, df $=$ N-1

TABLE 2
Gain in Shooting Score and Percentage of Hits in Archery

	Group I (two-day-a-week)			Group II (four-day-a-week)		
	Total	Men	Women	Total	Men	Women
Number	35	15	20	28	18	10
Score						
Range	15.40	14.50	14.70	18.70	18.70	13.70
Mean	3.77	3.48	3.98	6.14	6.21	6.02
SD	3.74	4.25	3.49	4.37	4.97	2.97
t	5.89*	3.05*	4.98*	7.31*	5.13*	6.08*
Hits						
Range	75	65	49	56	56	43
Mean	9.96	4.47	13.90	18.00	16.44	20.80
SD	13.99	14.92	11.72	14.30	15.35	11.66
t	4.11*	1.12	5.17*	6.55*	4.41*	5.59*

*Significant at the 1% level of confidence.

dence (Table 2). The gain in hits made by the men in Group I was not sufficient to be significant at the 5-per-cent level.

Using the mean gain for each group, the *t* test[3] for the significance of the difference of means was made. The *t* obtained for the difference of the scores between Group I and Group II, and the *t* for the difference in hits (Table 3) indicated that the difference in the gain of the two groups was significant at the 5-per-cent level of confidence for both scores and hits. This difference in gain in both instances was in favor of Group II—the four-day-a-week group.

In learning curves constructed for the men and women in each group, it was found that the four-day-a-week group showed greater total gain, as well as a more consistent gain throughout the total period. The men of Group I showed a more erratic learning rate, and the women of Group II a much sharper and more continuous rise after the eighth lesson. The learning curves for women, with respect to hits, showed a slightly more erratic rise for Group II than was indicated by the curves based on total score. For some unidentified reason, the learning curve for hits for men in the two-day-a-week group showed a sharp drop for the last three practices.

Table 3 indicates that there was no significant difference in gain in either scores or hits for men or women, between or within groups. What small differences there were favored Group II. This follows the findings in the over-all comparison of the two groups. A *t* significant at the 5-per-cent level was obtained for the difference in mean gain in hits between the men in the two groups.

There were no significant *t*'s obtained between Groups I and II for students of high motor ability or of low motor ability. A comparison of gains

[3] $t = \dfrac{M_1 - M_2}{\sqrt{\left(\dfrac{\Sigma x_1^2 + \Sigma x_2^2}{N_1 + N_2 - 2}\right)\left(\dfrac{N_1 + N_2}{N_1 N_2}\right)}}$, $df = N_1 + N_2 - 2$ (5, *p. 228*)

TABLE 3
t-Test of the Significance of the Difference of Mean Gain in Archery Within and Between Groups

Difference in Mean Gain	Shooting Score		Percentage of Hits	
	t	Level of Confidence	*t*	Level of Confidence
		(per cent)		*(per cent)*
II − I	2.28	5	2.25	5
II(M) − I(M)	1.63	10	2.19	5
II(W) − I(W)	1.53	20	1.41	20
I(M) − I(W)	.36*	80	2.10*	10
II(M) − II(W)	.11*	95	.74*	50
II(hiMA) − I(hiMA)	1.68	20	2.01	10
II(loMA) − I(loMA)	1.23	30	.31	80
I(hiMA) − I(loMA)	.17	90	.90	40
II(hiMA) − II(loMA)	.46	70	1.29	30

*difference favors women

Explanation of Code

I—Group I (two-day-a-week)
II—Group II (four-day-a-week)
M—Men
W—Women
hiMA—High motor ability
loMA—Low motor ability

between high and low motor ability within groups also failed to show any *t* large enough to be significant at the 5-per-cent level (Table 3).

It would appear from the data presented on archery that learning was more rapid, as measured by gain in mean score and percent of hits, when classes met four days a week rather than twice a week. For the subjects used in this study, there was apparently no significant difference in the amount of gain due to difference in sex or level of motor ability.

BADMINTON

In badminton three learning curves were drawn based on the scores from the three tests which were given at intervals throughout the practices. As it was not practical to test the students during every class period, the learning curves were of necessity based on fewer test scores than were those in archery. The wall volley test was given seven times and offers the most comprehensive picture of the learning rate in badminton. The high clear, because of the difficulties in administration, was given only three times and consequently resulted in a smoother learning curve than the others. The short serve test was given five times and gave nearly as good picture of progress as did the wall volley. In all three curves, there was a sharp rise from the first to the second testing and the curve did not rise again.

From a survey of these graphic presentations (see Figure II), it was found that there was very little real difference in the total pattern of the rate of learning between the two-day-a-week and the four-day-a-week groups with respect to the short serve and the high clear. In the wall volley, although the two groups began at approximately the same level of ability relative to the initial scores on this test, the two-day-a-week group made greater and more rapid gain than did the four-day-a-week group.

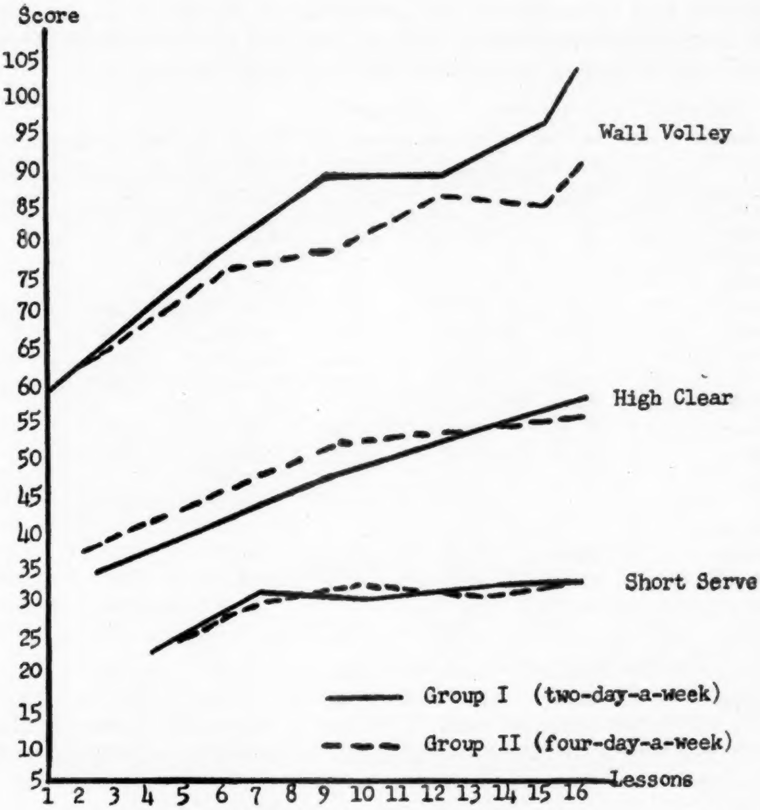


FIGURE II. Badminton Learning Curves

As in archery, the amount of gain in learning was determined by obtaining the difference between initial and final achievement scores. In all three of the skill tests, the initial evaluation used for determining the mean gain was the score made on the first administration of the test. For the final achievement, an average of the scores of the last two testings of the wall volley and short serve were used. In the high clear test, the scores on the last test only were used as the final evaluation. Table 4 presents the mean gain in all three of the skills tested for both groups plus the mean gain as measured by a composite T-score based on all three tests.

Table 4 gives the values obtained by using the *t* test of the significance of the difference of means between initial and final scores for the comopsite T-scores of the three tests, the wall volley test, the short serve test, and the high clear test. The *t*'s obtained indicated that with but two exceptions the differences between the initial and final scores were significant at the 1- or

2-per-cent level of confidence. The gain made by the men in Group I in the high clear was not significant at the 5-per-cent level and the women in Group I made a gain significant at the 5-per-cent level on the short serve.

TABLE 4

Amount of Gain in Wall Volley, Short Serve, and High Clear Tests in Badminton

	Group I (two-day-a-week)			Group II (four-day-a-week)		
Test	Total	Men	Women	Total	Men	Women
Number	35	12	23	41	17	24
Composite T-scores						
Range	26.51	23.83	24.62	21.33	17.67	21.33
Mean	9.68	9.69	9.67	7.84	6.94	8.48
SD	6.33	2.33	3.11	4.96	2.65	2.90
t	8.72*	14.47*	14.65*	10.60*	10.62*	14.05*
Wall Volley						
Range	85.50	79.00	71.50	64.00	60.00	46.00
Mean	40.03	36.42	41.91	27.11	24.59	28.90
SD	20.96	37.34	16.41	16.59	17.72	15.49
t	11.13*	3.24*	11.98*	10.31*	5.55*	8.95*
Short Serve						
Range	66.50	52.00	66.50	61.00	45.00	61.00
Mean	8.81	13.83	6.20	8.13	8.18	8.10
SD	14.44	16.99	12.10	13.62	11.59	14.89
t	3.55*	2.70†	2.40**	3.76*	2.83†	2.62†
High Clear						
Range	94.00	94.00	68.00	90.00	60.00	90.00
Mean	23.14	14.25	27.78	17.66	14.18	20.13
SD	24.82	25.31	23.25	21.12	19.99	21.53
t	5.43*	1.87	5.60*	5.29*	2.84†	4.48*

*Significant at the 1% level of confidence.
†Significant at the 2% level of confidence.
**Significant at the 5% level of confidence.

When the t test was applied to the data from Table 4, it was found that the obtained t for the wall volley was significant at the 1-per-cent level (Table 5). A t was obtained for the short serve which was significant at the 5-per-cent level, while the t for the high clear was too small to be significant at the 5-per-cent level. All three of these differences were in favor of the two-day-a-week group. The initial and final scores for the wall volley, short serve, and high clear were T-scored and these T-scores were combined into a composite score. In testing the significance of the difference btween the mean gain of the two groups as measured by this composite T-score, a t was obtained which was significant at the 1-per-cent level, with the difference in gain in favor of Group I.

The learning curves for men and women for the wall volley, short serve, and high clear all showed a slightly more accelerated rate of learning for the two-day-a-week group than for the four-day-a-week group. This difference was most marked for women in the wall volley test. The other differences appeared to be very slight or non-existent. The results of the t test of

TABLE 5

t-Test of the Significance of the Difference of Mean Gain in Badminton Within and Between Groups

Difference in Mean Gain	Composite T-scores		Wall Volley		Short Serve		High Clear	
	t	Level of Confidence	*t*	Level of Confidence	*t*	Level of Confidence	*t*	Level of Confidence
		(per cent)		*(per cent)*		*(per cent)*		*(per cent)*
I – II	3.45	1	2.96	1	2.08	5	1.03	40
I(M) – II(M)	1.29	30	1.11	30	1.03	40	.01	99
I(W) – II(W)	.69	50	2.74	1	.47	90	1.15	30
I(M) – I(W)	.01	99	.58*	60	1.27	20	1.52*	20
II(M) – II(W)	.89*	50	.81*	50	.02	99	.88*	50
I(hiMA) – II(hiMA)	.53	60	.85	50	1.58	20	.18§	90
I(loMA) – II(loMA)	.20	90	2.10	10	.98§	40	.41	70
I(hiMA) – I(loMA)	.02†	99	1.08†	30	.73	50	.06	99
II(hiMA) – II(loMA)	.46†	70	.19	90	1.81†	10	.84	50

*difference favors women
§difference favors four-day-a-week group
†difference favors low motor ability

Explanation of Code

I—Group I (two-day-a-week)
II—Group II (four-day-a-week)
M—Men
W—Women
hiMA—High motor ability
loMA—Low motor ability

the difference substantiated this observation. The *t* obtained for the difference between the mean gain of Group I and Group II for women was significant at the 1-per-cent level of confidence and in favor of the two-day-a-week group. The rest of the *t* values were not of sufficient magnitude to be significant at the 5-per-cent level. In testing the difference between the gain made by men and women within each group, the *t* values obtained were too small to be significant at the 5-per-cent level, but indicated slightly greater gain for women than for men.

In comparing the rates of learning in the three skill tests between Group I and Group II for students with high motor ability and for students with low motor ability, it was found that, although the character of the curves differed, they showed no essential differences in learning rates. The one exception to this was found in the comparison of the mean gain of those subjects with low motor ability in Group I with those in Group II. However, the *t* obtained was only large enough to be significant at the 10-per-cent level. The other *t* values obtained for high or low motor ability did not approach even the 10-per-cent level of confidence. In Group II a *t* significant at the 10-per-cent level and in favor of the low motor ability group was obtained. This was the only difference which even approached this level of significance.

From a survey of the statistical analysis of the data, it would appear that the two-day-a-week distribution of practice had a slightly more favorable effect on the rate of learning of college students in badminton, as measured by the composite T-scores of the three tests, or by wall volley alone, than did the four-day-a-week distribution. Apparently neither the level of motor ability nor sex had any effect on this rate of learning, as was also true in archery.

It is not within the scope of this study to attempt to determine why learning

in archery appeared to be favored more by the four-day-a-week distribution of classes while learning in badminton was favored more by the two-day-a-week pattern. Other experimenters have found that, in certain learning situations, learning in the early stages seemed to be improved when the practice periods were massed, but that, once fundamental learning has taken place, learning was more rapid if the practices were distributed over longer periods of time. Might it be possible that as badminton is only one of several racket games that previous experience in these other games might affect learning in badminton? Archery, on the other hand, is an activity quite unique in the skills necessary for success, therefore the students would not have been apt to have much experience which might help them in learning archery.

Is the difference of the effect of relative spacings of practices on learning in archery and badminton due, on the other hand, to a difference in the complexity of the skills of the activities concerned? Archery is an activity in which it is necessary to learn relatively few fundamental skills, but in which very slight deviations from "perfection" in shooting form may markedly affect the success of the activity. In badminton, although there are many skills to be learned, a certain degree of success and pleasure may be attained the first time a student takes a racket in hand, without the necessity of first perfecting the requisite skills. Whatever the reason, there seems to be little doubt that th difference lies, at least in part in the activities themselves and not entirely in the way in which the practices were distributed.

Summary and Conclusions

It was the purpose of this study to compare the effectiveness of learning as measured by the mean gain from initial to final evaluations in archery and badminton classes taught under either a two-day-a-week or a four-day-a-week distribution of class periods. Regular college physical education classes were used in the desire to make the experimental situation more practical and more nearly comparable to the average conditions under which physical education classes are taught.

The preponderance of evidence from past experimental studies has favored distributive over massed practice for more effective learning. In view of past experimental set-ups, the present study might more appropriately have been called a comparison of two types of distributive practice on learning rate, as the four-day-a-week pattern hardly approaches the massed practice patterns usually used.

The results of the study seem to indicate certain trends:

1. In archery the four-day-a-week distribution of practices seemed to result in more rapid learning. However, in badminton more effective learning appeared to be facilitated by the two-day-a-week distribution.

2. These results were not affected by either sex or level of motor ability. However, the women did show slightly greater gain than did the men. Also the students with high motor ability tended to show slightly more gain than those with low motor ability.

The writer feels that, although the results of this study are inconclusive, the study of the effect of distribution of practices on learning in physical education is a field which merits more study. It would be of interest to compare the results of this study with similiar ones done at the high-school and junior-high-school levels, and with other activities at both college and secondary levels.

REFERENCES

1. Ammons, Robert B. Acquisition of Motor Skill: Effects of Initially Distributed Practice on Rotary Pursuit Performance, *Journal of Experimental Psychology.* **40:** 777-87 (1950).
2. Ammons, Robert B., Effect of Distribution of Practice on Rotary Pursuit Hits, *Journal of Experimetnal Psychology,* **41:** 17-22 (1951).
3. Cook, Barbara S., and Ernest R. Hilgard, Distributed Practice in Motor Learning; Progressively Increasing and Decreasing Rests, *Journal of Experimental Psychology,* **39:** 169-72 (1949).
4. Dore, Leon R., and Ernest R. Hilgard, Spaced Practice and Maturation Hypothesis, *Journal of Psychology,* **4:** (1937).
5. ———, Spaced Practice as a Test of Snoddy's Two Processes of Mental Growth, *Journal of Experimental Psychology,* **23:** 359-74 (1938).
6. Duncan, Carl P., The Effect of Unequal Amounts of Practice on Motor Learning Before and After Rest, *Journal of Experimental Psychology,* **42:** 257-64 (1951).
7. Guilford, J. P., *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill Book Co., Inc., Second Edition, 633 pp. (1950).
8. Harmon, John M., and Arthur G. Miller, Time Patterns in Motor Learning, *Research Quarterly,* **21:** 182-7 (1950).
9. Hilgard, Ernest R., and M. Brewster Smith, Distributed Practice in Motor Learning; Score Changes Within and Between Daily Sessions, *Journal of Experimental Psychology,* **30:** 136-46 (1942).
10. Knapp, Clyde G., and W. Robert Dixon, Learning to Juggle: A Study to Determine the Effect of Two Different Distributions of Practice in Learning Efficiency, *Research Quarterly,* **21:** 331-6. (1950).
11. Kuhl, Louise, *A Study of the Learning Rate of College Women in Certain Physical Activities as Related to Their General Motor Ability,* Unpublished Ph.D. thesis, State University of Iowa (1941).
12. Lashley, K. S., *The Acquisition of Motor Skills in Archery,* Paper from the Department of Marine Biology, the Carnegie Institute of Washington, **7:** (1915).
13. Nance, R. D., Effects of Pacing and Distribution on Intercorrelations of Motor Abilities, *Journal of Experimental Psychology,* **37:** 459-72 (1947)
14. Scott, M. Gladys, and Esther French, *Evaluation in Physical Education,* St. Louis: The C. V. Mosby Co, 348 pp. (1950).
15. Snell, Catherine C., *A Study of Rates of Learning in Selected Sports as Related to General Motor Ability,* Unpublished Ph.D. thesis, State University of Iowa (1948).
16. Snoddy, George S, *Evidence for Two Opposed Processes in Mental Growth,* The Science Press Printing Co. (1935).
17. Travis, Roland C., Length of the Practice Period and Efficiency in Motor Learning, *Journal of Experimental Psychology,* **24:** (1939).
18. ———, Practice and Rest Periods in Motor Learning, *Journal of Psychology,* **3:** (1937).

(Submitted 7-27-53)

Research Abstracts

Prepared by the Research Abstracts Committee of the National Council of the Research Section, PAUL A. HUNSICKER, Chairman

Anatomy

19. Brozek, Joseph. Measuring nutriture. *American Journal of Phys. Anthropology,* **11:** 2 (June 1953).

Objective criteria for the assessment of nutriture are needed to evaluate the status of different populations. The anthropologist is concerned with evaluating over- and under-nutrition, not with malnutrition. "Body build" may be broken down into (1) body composition (mass of principal tissues, or in terms of chemically defined constituents); (2) body structure (tissue composition of body segments, especially limbs—the "internal body form"); and (3) the external body form—the anthropologist's principal concern up to the present. The need for a practical solution for the use of anthropometric measurements in the evaluation of nutritional status represents a challenge to the physical anthropologist.—*Frank D. Sills.*

20. Evans, F. Gaynor. Methods of studying the biomechanical significance of bone form. *American Journal of Physical Anthropology,* **11:** 3 (Sept. 1953).

Methods used to study stresses and strains in bones may be divided into (1) mathematical analysis of the stress and mechanical behavior of sections of bones, (2) the study of stresses and strains in models of bones, and (3) the study of stresses and strains produced in intact bones. For the second and third methods, recently developed engineering techniques have been employed.

The method of mathematical analysis may be criticized for (1) the assumption is made that the bone is composed of uniform homogeneous material (which is not true), (2) the results are in two instead of three dimensions (as is the actual case), and (3) the formulae used are those derived from similar studies on other substances such as steel (which may not apply). The study of models is also unsatisfactory, primarily because of inability of scientists to accurately reproduce the characteristics of bone.

In the study of intact bone, the application of a lacquer which cracks under stress has proved valuable. It is believed that the use of this "Stresscoat" lacquer, followed by the use of strain gauges, should result in accurate measurement of the magnitude of strain in a local area. Variations in the physical properties of a given bone will, of course, influence the study of its stresses and strains (including such factors as age, state of health, nutrition, etc.).—*Frank D. Sills.*

21. Ferguson, Albert Barnett. A study of the action at varying lengths of two muscles of opposing functions, tibialis anterior and gastrocnemius in the rat. *Jour. of Bone and Joint Surg.,* **35-A:** 3 (July 1953).

The two muscles were tested by means of successive isometric contractions at steadily increasing lengths. A Grass stimulator was used which put out a square-wave stimulus of one-millisecond duration. A strain gauge was used to record the tension on the muscles.

The developed tension of the tibialis always rose to a maximum on the left side of the resting-tension curve, while the gastrocnemius developed increasing heights of tension slowly, so that the maximum was reached to the right of the resting-tension curve. In summary, the tibialis performed at maximum just beyond its initial resting length, while the gastrocnemius worked at maximum, after its resting length had been increased a greater distance. The author concluded that there is no standard configuration applicable to length-tension diagrams and that the point of maximum developed tension will vary with structure and function.—*Frank D. Sills.*

22. RAY, ROBERT D., ET AL. Growth and differentiation of the skeleton in thyroidectomized-hypophysectomized rats treated with thyroxin, growth hormone and the combination. *Jour. of Bone and Joint Surg.*, **36-A**: 1 (1954).

An effort was made to determine the influence of thyroxin and of pituitary growth hormone on skeletal growth and maturation (as determined by the appearance of the centers of ossification) and their subsequent fusion-boneage). A study of the tibia and humerus, following the 60-day experiment period, revealed that growth and skeletal maturation were equally retarded following the combination of thyroidectomy and hypophysectomy. Thyroxin injections produced marked maturation and only moderate growth, while injections of the pituitary growth hormone produced a definite increase in body weight and skeletal dimensions, but did not produce skeletal maturation. The administration of both hormones to the animals, following both operations, restored the coordination between skeletal growth and maturation.—*Frank D. Sills.*

Education

23. DREGER, RALPH MASON. A simple course evaluation scale. *J. of Experimental Educ.*, **22**: 2, pp. 145-153 (Dec. 1953).

A simple to use and easy of access course evaluation scale was presented. Its quantitative uses are to show the direction of attitude of classes, chronological trends, and possibly group comparisons.—*D. B. Van Dalen.*

24. RYANS, DAVID G. A statistical analysis of certain educational viewpoints held by teachers. *J. of Experimental Educ.*, **22**: 2, pp. 119-131 (Dec. 1953).

Two forms of an Educational Viewpoints Inquiry were constructed and administered respectively to 551 elementary and secondary teachers in geographically scattered colleges of teacher education. The instrument was a short form made up of 20 items, each item forcing a choice between contrasting viewpoints regarding educational practices. The inquiry supports the belief that teachers' educational viewpoints are not highly systematized or organized. They tend to be organized into relatively independent clusters with respect to curricular organization, pupil participation in class planning, academic achievement standards, etc.—*D. B. Van Dalen.*

25. WILKINS, W. D., and LUCY GROSS. Usefulness of educational periodicals for research. *School and Society*, **79**: 2024, pp. 9-11 (Jan. 9, 1954).

Periodicals cited most frequently for research purposes were appraised. The *Review of Educational Research*, the *Journal of Educational Research*, and the *Education Digests* over a period of five years, 1948-52 inclusive, were employed to determine frequency of citation in critical reviews or abstracts.—*D. B. Van Dalen.*

Health

26. BEDOIAN, V. H. Mental health analysis of socially over-accepted, socially under-accepted, over-age and under-age pupils in the sixth grade. *J. of Educ. Psych.*, **44**: 6, pp. 336-371 (Oct. 1953).

The Thorpe, Clark, Tiegs' *Mental Health Analysis* was administered to 743 sixth-grade Los Angeles district pupils. Nine months below or above the mean chronological age of their half-grade placement determined the age placement. Data indicated that under-age and at-age pupils earned significantly better mental health scores than the average pupils. There was no significant difference between under-age and at-age pupils. The socially over-accepted produced significantly higher mental health scores than the under-accepted pupils. Superior economic status group had better mental health than pupils who were ignored, unwanted, and disliked by their peers. The "stars of attraction" excelled in better mental health.—*D. B. Van Dalen.*

27. Committee on Teacher Preparation in Health Education, American School Health Association. Recruitment and placement of health education majors. *J. School Health,* **24:** 1 (Jan. 1954).

A questionnaire study surveyed institutions offering an undergraduate and/or graduate health education major, but not having a School of Public Health. Replies were secured from 32 of 35 colleges and universities offering undergraduate health education majors. In a few their programs are still largely paper programs.

Nineteen of the 32 offer graduate health programs. Placement of graduates with a Bachelor's degree has been difficult; it was less difficult to place those with Master's degrees, and relatively easy to place graduates with doctoral degrees.

The committee feels there is real need (1) to stimulate more and better research in health education; (2) to develop more advanced graduate level health education courses, especially for doctoral students; and (3) to develop in a few institutions, top-flight graduate health education programs leading to the research type of doctoral degree (Ph.D.)—*John H. Shaw.*

28. DeKruif, Paul. Bursitis. *Today's Health,* **32:** 24-25, 52-53 (Feb. 1954).

Painful shoulder, bursitis, is thought of as a minor malady, but millions suffer its torture and annoyance every day. The soreness which results is due to tendon inflammation and calcium deposits. The strapping of arms down to avoid movement is heartily condemned. As the red-hot pain of acute bursitis subsides, the patient should bend over with the arms relaxed and gently swing the arm like a pendulum, from front to rear, side to side, and round and round. The use of hormones (ACTH, cortisone, or hydrocortisone) injected directly into the tender spots will do much to eliminate the pain and discomfort of red-hot bursitis.—*J. Grove Wolf.*

29. DeWeese, A. O., M.D. Teacher's Health Inspection, *J. School Health,* **23:** 8 (Oct. 1953).

Dr. DeWeese suggests it is time to give the classroom teacher more definite and more scientific procedures in helping her play her part in the physical inspection and health appraisal of our children. He outlines procedures for checking the following: 1. Muscle Tone and Graceful Walking, 2. Enlarged Cervical Glands, 3. Nose and Mouth Breathing, 4. Examination of the Eyes, 5. The Throat and Tongue, 6. The Tongue, 7. Teeth and Gums, 8. Condition of the Scalp, 9. Nutrition, 10. Thyroid Gland and Ears, 11. Cleanliness, Nutrition, and Skin Infections, 12. When To Use The Clinical Thermometer. These procedures have been used with satisfactory results.—*John H. Shaw.*

30. Johnson, Lucille, R.N. Hyperopia in school children: Its detection with a plus sphere lens, and its significance in reading. *J. School Health,* **23:** 9 (Nov. 1953).

This is a report of screening with a plus 1.75 lens in addition to the Snellen test already employed in Oakland. The developers of this test, the Massachusetts Department of Public Health, reported about 7 per cent of all school children will test farsighted. The Oakland study showed an average of 1.5 per cent and in no age group did it come near 7 per cent. It also reported a decreasing percentage of hyperopia among older children which tends to hear out the theory that as the eyeball grows, hyperopia will decrease.—*John H. Shaw.*

31. Millman, Max. How safe are reducing diets? *Today's Health,* **31:** 27, 53-58 (Oct. 1953).

After indicating several objections to ready-made diets, Millman recommends several well-founded nutritional rules for the construction of a sound reducing diet. The diet must be low in calories, low in fat, moderately low in starches or sugars, fairly high in protein, adequate in water, minerals and vitamins, ample in bulk. Remember there is no food or drink whose caloric value is lower than zero and, advertising claims notwithstanding, there is no patent medicine on the market capable of reducing a single ounce of weight without risk. The only safe and effective reducing diet is one based on sound, scientific medical knowledge.—*J. Grove Wolf.*

32. Wolfe, Burton H., The cured cancer club. *Today's Health*, **31**: 24-26 (Dec. 1953).

The earlier cancer is detected, the better your chances for a cure. Danger signals in cancer, which, if they appear, should mean an early visit to your physician, include: (1) Any sore that does not heal; (2) A lump or thickening in the breast or elsewhere; (3) Unusual bleeding or discharge; (4) Any change in a wart or mole; (5) Persistent indigestion or difficulty in swallowing; (6) Persistent hoarseness or cough; and (7) Any in bowel habits. The Cured Cancer Club of Washington (1415 Eye Street, N.W., Washington 5, D. C.) is available to help by answering your questions and providing information to assist those afflicted with this disease.—*J. Grove Wolf.*

Physiology

33. Sclar, Meyer. *Heart and Circulation,* New York: Froben Press, 1953, p. 357.

The book is composed of two sections; namely, a section on diseases of the heart and one on diseases of the vascular tree. The basic concepts of diagnosis and treatment of cardiac disorders are presented in condensed form and the text should prove to be a useful primer for the practicing cardiologist.

The author gives a brief background of the embryology, anatomy, and physiology of the heart and circulatory system. Several chapters are devoted to methods of diagnosing cardiac disorders. The most common diseases are covered separately and the etiology, symptoms, diagnosis, and treatment are given.

The terse treatment of the material plus the omission of footnotes and bibliography permits the author an opportunity to present a considerable amount of factual information in a small volume.—*Paul Hunsicker.*

Recreation

34. Anderson, Jackson M. Leadership: the most important quality in professional recreation management. *Industrial Sports,* **14**: 25, 28, 33 (Oct. 15, 1953).

A questionnaire was sent to 100 member companies of the National Industrial Recreation Association in an attempt to develop personnel standards for the position of director of recreation in business and industrial firms.

The findings of the study indicated that the director of recreation should be a college graduate with industrial recreation as the major field of study and physical education as the minor field, or should have a substantial amount of practical experience in the field of recreation. In the educational background of the director, preparation in the following subjects was considered desirable: industrial recreation, physical education, industrial relations, psychology, sociology, English, public speaking, sports activities, and nature study.

The following types of previous experience were considered desirable in qualifying one for the position of director: supervisor of recreation, participant in varsity sports, physical education instructor, community recreation director, athletic coach, athletic director, personnel manager, YMCA or YWCA physical director, camp director, and scout leader. It was felt that the director should possess the following qualifications and characteristics: (1) leadership ability, (2) pleasing personality, (3) thorough knowledge of recreation, (4) good judgment, (5) sociability, (6) ability as a public speaker, (7) initiative, (8) knowledge of budgets and accounting, (9) neat appearance, and (10) good health.—*Jackson M. Anderson.*

35. Aurand, Rex. Bowling in Junior High School. *Athletic Journal,* XXXIV: 13, 61-62 (Nov. 1953).

The author has used bowling as a regular physical education class activity since 1948. The school has no regular alleys; therefore the gymnasium was used. The floor was protected with rubber mat runners, gymnasium mats were used for pin and ball deflections, and locker room benches were placed in a "V" shape between the alleys. Pin setters were students in the class. Safety precautions, such as using the foot instead of the hand when the ball is rolled back, were agreed upon. A little ingenuity can accomplish much in the way of added program opportunities.—*J. Grove Wolf.*

36. Fitzgerald, Gerald B. and Albert Meuli. Supply and placement of recreation graduates. *Quarterly Bulletin of the American Recreation Society*, **5**:8, 10-12 (July 1953).

The study included reports from 43 colleges and universities. These institutions represented 90% of those offering recreation degrees in 1951-52. Twenty of the reporting institutions offered graduate degrees in recreation, enrolling a total of 260 graduate students in this field. The total enrollment of undergraduate and graduate recreation majors was 1,531. Approximately 43% of those majoring in recreation were women. Undergraduate degrees were awarded to 374 and graduate degrees to 138 persons during the year. Data were submitted on job placement for 353 (70%) of the 512 degree recipients. Of the 353, 71% were placed in full-time recreation positions, 13% in part-time positions in recreation, and 16% in positions outside the field of recreation.

The types of recreation positions and the percentage of graduates placed in each were as follows: public community recreation, 47.7%; private community recreation, 19.6%; hospital recreation, 12.3%; institutional recreation, 4.8%; industrial recreation, 3.9%; state-level agencies, 2.6%: college teaching in recreation, 2.6%; armed forces civilian recreation positions, 2%; commercial recreation, 2%; church recreation, 1.6%; college student union positions, 0.6%; and year-round camping positions, 0.3%.

After comparing the findings with those of four earlier studies, the following trends were pointed out: the number of institutions granting degrees in recreation continues to rise on the undergraduate level; the college recreation department and the college placement bureau were the agencies mainly effective in placing degree recipients in positions.—*Jackson M. Anderson.*

37. National Recreation Assn. Dust elimination. *Recreation*, **46-A**: 94-95 (May 1953).

Dust elimination was one of the three topics on which information was requested in an inquiry on surfacing recreation areas, conducted in 1952 by the national committee studying this subject. The findings of the survey were based on an analysis of the 153 reports received from park and recreation executives.

Some findings of the study were as follows: (1) The greatest dust problem appeared to be centered on ball diamond infields and, in lesser degree, on tennis courts, apparatus areas, and bridle trails; (2) Water was the most widely used agent in the attempt to eliminate dust, being reported by 51% of those surveyed; (3) Water was the most costly agent to apply and perhaps to control, unless the user had sufficient funds to install a modern sprinkler system; (4) Approximately 19% of the executives submitted information on the use of oil, with 12 areas reporting successful control, seven areas mediocre results, and ten areas unsatisfactory experiences; (5) The failure of the oil manufacturer to furnish complete use specifications for his product for the various types of soils was the major reason for many unsuccessful uses; (6) Several areas using oil reported clothing-stain experiences, messy soil conditions for several days, and the tracking of oil into adjacent buildings; and (7) Calcium chloride was used successfully in many areas with high humidity, but it was found to be ineffective in most areas with low humidity.—*Jackson M. Anderson.*

Miscellaneous

38. Reed, Harold J. An investigation of the relationship between teaching effectiveness and the teacher's attitude of acceptance. *J. of Experimental Educ.*, **21**: 4, pp. 277-325 (June 1953).

Three secondary schools and 160 teachers were chosen to participate in the study. A relationship far beyond chance expectancy was found to exist between teacher's effectiveness in the classroom as evaluated by the students and those aspects of a teacher's personality, organization, or attitude, which permits him to be an acceptable person. It was not possible to predict safely which teachers the administrator would judge to be effective. The Sentence Completion Test correctly identified the effective and ineffective teachers as rated by students in better than 75% of the cases—*D. B. Van Dalen.*

Guide to Authors

IN LINE WITH the over-all goal of making Association publications yield the greatest value to the individual and the profession, the following is a guide for the preparation of research manuscripts. The information below recognizes general techniques being employed by research publications similar to the *Research Quarterly*. When copy is prepared in accordance with these instructions, all Association research studies will follow a standard style.

Manuscripts

Manuscripts should be sent to the Editor (AAHPER, 1201 Sixteenth Street, Northwest, Washington 6, D. C.), who will see that each one is read by at least three members of the *Research Quarterly* Board of Associate Editors. On the basis of the three reviews, the Editor will advise the author as to the suitability of the paper or the desirability for revision. Papers are not judged by arbitrary standards but on their content of new research results in the field of physical education, health education, and recreation, presented with the greatest brevity compatible with scientific accuracy and clarity (see October 1951 *Quarterly*, pp. 392-4).

Since three members of the Board of Associate Editors review an article, it is requested that three clear copies of the manuscript be submitted in order to facilitate reviewing. A fourth copy of the article should be retained by the author. Only one copy of any charts, photographs, drawings, graphs, or similar illustrative material need be submitted. However, since such material must be sent to each reviewer in turn, more time must be allowed for the reviews.

Typewritten manuscript should be double-spaced on white paper of ordinary weight and standard size ($8\frac{1}{2}$ x 11 inches).

The sheets of manuscript should be kept flat and fastened with clips which can be removed easily. The pages of the typewritten copy should be numbered consecutively in the upper right-hand corner.

Paragraphs should be numbered consecutively throughout the manuscript to facilitate ease of reference in case of revision.

Headings

The article should be arranged so as to indicate relative values of heading and subheadings.

Usually four gradations are sufficient: (a) article title, (b) first subhead appearing in boldface aligned left on page (underscored in manuscript with wavy line) (c) second subhead (if necessary) appearing in small caps aligned left on page, (d) third subhead, to appear in italic (underscored in manuscript), not centered, but run in at the beginning of the paragraph or section.

All headings should be typed in lower case with initial capitals, except for (c) above, which should be typed in capital letters.

FOOTNOTES

Footnotes are not to be used for references or literature citations. They are rather used for the purpose of acknowledgment, special explanation, supplementary information, etc. (*See examples below.*)

Type footnotes (if any) on separate sheets, as many footnotes as convenient being written on a sheet. Footnotes should be numbered from 1 up for each article; a corresponding numeral appearing in the text. Asterisks should not be used.

Examples of Footnotes:

[1] This study was made under the direction of Dr. Arthur T. Slater-Hammel in the Research Laboratories, School of Health, Physical Education, and Recreation, Indiana University, Bloomington, Indiana.

[2] All measurements of the hand were recorded in centimeters and height was recorded in inches. The hand measurements were taken by Everett and reliability coefficients of above .90 were found for each measurement used in the study.

[3] For their wholehearted co-operation in facilitating collection of the data, special gratitude is extended to Superintendent Clarence Hines and the 1950-51 principals of the Adams, Condon, Edison, Francis Willare, Harris, Howard, Lincoln, River Road, and Whiteaker schools.

Documentation

CITATIONS OF LITERATURE

Citations of literature should be segregated alphabetically by author's last name at the end of each article, under the caption of "REFERENCES." *Do not treat them as footnotes.* (*See above.*)

The literature citations, listed alphabetically, should be numbered consecutively, their location in the text being indicated by corresponding numbers written in full size and enclosed in parentheses: for example, (1) (2, 3). If there are several references in the text to a citation, the specific pages may be indicated thus: (1, *p. 117*) (1, *pp. 162-3*).

A uniform style should be maintained in writing citations. Do not enclose titles of chapters and articles in quotation marks. Italicize (underscore in manuscript) names of books and periodicals, bulletins, etc. (*See examples below.*)

Uniform sequence of data should be observed, as follows: *For a book*—Author's name (last name first); title of article or chapter; name of book; place of publication; publisher; year date. *For a periodical*—Author's name (last name first); title of article or chapter; name of periodical; volume number; inclusive page numbers; year date.

Examples of References Appearing at End of Article:

1. American Association for Health, Physical Education, and Recreation. Suggested platforms for health education. *Journal of the American Association for Health-Physical Education-Recreation* **18**:436 (Sept. 1947).
2. American Association of School Administrators. *Health in Schools.* Revised edition. Washington, D. C.: the Association, a department of the National Education Association. pp. 266-7.

3. DEAVER, G. G., Exercise and Heart Disease. *Research Quarterly,* **26**: 24-34, 1939.
4. OGDEN, JEAN, AND JESS OGDEN. *Small Communities in Action.* New York: Harper & Brothers, 1946.
5. POTTER, JOHN NICHOLAS. *Physical Fitness of Junior High School Boys.* Unpublished Master's thesis, University of California, Berkeley, 1942

Tables

Each table should have a descriptive heading and should be specifically referred to in the text by number, *e. g.*, "Table 10," etc., never as "the above table" or "the following table." Number tables from 1 up for the entire manuscript, using Arabic numerals. Do not duplicate data by giving it in *both* tables and graphs.

Tables should be double-spaced typewritten, like the rest of the material in the manscript. They should be typed on separate sheets, as the printer will set them on a different machine from the one used for the text matter. If a table continues on a second sheet, it is not necessary to repeat the boxheads, since the printer will repeat from the original boxheads, when necessary.

The word "TABLE" should be written in capital letters, as: "TABLE 1"; the table title should be written in lower case letters with initial capitals, and centered over the table. Tables should be ruled as desired, except that no rules will appear at the extreme right and left edges of the table. No double rules are to be used, unless necessary for clarity.

Illustrations

Illustrative material is of two types: pen and ink drawings, which are reproduced by the line engraving process; and photographs, wash drawings, stipple drawings (in short, anything containing shading), which are reproduced by the halftone process.

Line engravings are always treated as text figures and should be so designated. All drawings should be made with India ink, preferably on white bristol board plate, 1 ply or 2 ply, which is sufficiently transparent to permit tracing if back lighting (e.g., a window pane) is used. Avoid graph paper for the reproduction copy, as the printing interferes with proper inking and the paper permits no corrections. Sometimes it is desirable to ink in the principal guide lines so that the curves can be more easily read.

Lettering should be plain and large enough to reproduce well when the drawing is reduced to the dimensions of the printed page (4⅛ x 7 inches). Most figures can be advantageously drawn for a linear reduction of one-half or one-third. Be sure to draw the lines heavy enough so that they will not be overly thin after reduction. Explanatory lettering should be included within the chart. Typewritten lettering does not reproduce well; it is much better to use a LeRoy or similar lettering device.

Care should be taken not to waste space, as this means greater reduction and a less satisfactory illustration. Often it is possible to combine several curves in one figure and thus not only save space but enable the reader to make comparisons at a glance.

Halftones are treated as figures and should be so designated. Frequently, several halftones can be grouped to form an attractive full page, in which case they should be numbered consecutively, in Roman numerals. Photographs should be in the form of clear black-and-white prints on glossy paper. Care should be taken to see that they cannot be bent or folded in handling *and paper clips should not be used.* All imperfections in the original copy are reproduced.

Figures should each be numbered consecutively from I up for the entire manuscript. Use Roman numerals to number figures, and Arabic numerals to number tables. The legends for the illustrations should be typed upon a separate sheet placed at the end of the manuscript. Care should be taken to indicate plainly in the text the exact location of all illustrations and tables.

The Association will assume complete engraving expense.

Special Points of Style

USE OF NUMBERS

Use Arabic figures for all definite weights, measurements, percentages, and degrees of temperature (for example: 2 kgm., 1 inch, 20.5 cc., 300°C.). Spell out all indefinite and approximate periods of time (for example: over one hundred years ago, about two-and-one-half hours). For numerals used in a general sense, spell out numbers through ten and use Arabic figures for 11 and over (seven times, five years old, 11 students).

ABBREVIATIONS AND SYMBOLS

Standard abbreviations should be used whenever the weights and measurements are used with figures, i.e., 10 kg., 6.25 cc., etc. The forms to be used (for both singular and plural) are: ft., ft.-lb., ft/sec, in., yd., min., hr., sq. ft., sq. in., rpm. *Gram* should be spelled out in all cases to avoid possible confusion with *grain;* also spell out *mile.* All obscure and ambiguous abbreviations should be avoided. Symbols used should follow the notation listed in *Research Methods* (AAHPER), pp. 518-20 and 522-25. The most common are given below:

M = mean	r = Pearson correlation
Mdn = median	r_{bis} = biserial correlation
N = number of individuals	r_{1I} = reliability coefficient
n = number of measurements	χ^2 = chi square
σ = standard deviation	F = variance ratio
σ_M = standard error of mean	t = Student (Fisher) t ratio

Per cent should be two words. Use per cent sign (%) in tables or when it appears in parentheses in text.

Proofreading

The author will receive his original manuscript and any engraver's proofs with the galley proofs of his article for correction. A reprint order blank will be enclosed for the author's convenience.

Corrected proofs and original manuscripts are to be returned within 48 hours by first-class mail to the Editor, AAHPER, 1201 Sixteenth Street, Northwest, Washington 6, D. C.